



12

EUROPEAN PATENT APPLICATION

21 Application number: 94111298.9

51 Int. Cl. 6: C12Q 1/02, C12N 9/02,
C12N 15/53, C12N 15/81,
//C12Q1/26

22 Date of filing: 20.07.94

30 Priority: 20.07.93 JP 201120/93
21.07.93 JP 180246/93
30.07.93 JP 208279/93

43 Date of publication of application:
22.03.95 Bulletin 95/12

34 Designated Contracting States:
CH DE FR GB LI

71 Applicant: SUMITOMO CHEMICAL COMPANY,
LIMITED
Kitahama 4-chome 5-33
Chuo-ku
Osaka 541 (JP)

72 Inventor: Hayashi, Koji
14-7, Mefu 2-chome
Takarazuka-shi,
Hyogo-ken (JP)
Inventor: Sakaki, Toshiyuki
1-30-68, Higashitoyonaka-cho
Toyonaka-shi,
Osaka-fu (JP)

Inventor: Yabusaki, Yoshiyasu
6-2-1, Minatojimanakamachi,
Chuo-ku
Kobe-shi,

Hyogo-ken (JP)
Inventor: Komai, Koichiro
4-9-17, Sakuragaoka
Minoo-shi,

Osaka-fu (JP)
Inventor: Kaneko, Hideo
63, Koyamaitakura-cho,
Kita-ku

Kyoto-shi,
Kyoto-fu (JP)
Inventor: Nakatsuka, Iwao
3-8-14, Hinomine,
Kita-ku
Kobe-shi,
Hyogo-ken (JP)

74 Representative: VOSSIUS & PARTNER
Postfach 86 07 67
D-81634 München (DE)

54 Method for safety evaluation of chemical compound using recombinant yeast expressing human cytochrome P450.

57 There is disclosed a method for evaluation of the safety of a chemical compound, which includes the steps of: (a) reacting a chemical compound with recombinant yeast cells expressing, or in other words producing, human cytochrome P450 molecular species P450 1A2, P450 2C9, P450 2E1 and P450 3A4 together with a yeast NADPH-P450 reductase, which may be in the form of a fused enzyme with each of said human cytochrome P450 molecular species, or with the cell free extracts of the yeast cells; and (b) analyzing the resulting metabolite to determine the safety of the compound. According to this method, it can be determined whether a test compound will be converted into a carcinogenic or mutagenic form through the metabolism in the human liver, and whether the test compound or its metabolite has mutagenicity.

EP 0 644 267 A2

The present invention relates to a method for evaluation of the safety of a chemical compound using recombinant yeasts expressing human cytochrome P450.

The cytochrome P450 is an enzyme catalyzing the mono-oxygenation of a substance in the human liver.

5 It is known that recombinant human cells expressing heterogeneous human cytochrome P450 species have been used for determination of metabolisms and toxicities of chemical substances. However, this method is unsatisfactory as a method of evaluation of the safety of chemical compounds partly because the kinds of the human cytochrome P450 species expressed by the cells and the levels of the expression are so limited that the amount of metabolite obtained is not enough for determination of the metabolism and
10 toxicity, and partly because it requires not only a high density culture technique but a high cultivation cost. Accordingly, there has been a great demand for developing an advantageous method.

As a result of the extensive study, the present inventors have found that yeasts are particularly suitable as hosts for production of human cytochrome P450 and yeast NADPH-P450 reductase to be used in in vitro determination of metabolisms and toxicities of chemical substances because yeasts grow so rapidly and
15 can stably express both the human cytochrome P450 and yeast NADPH-P450 reductase at high expression levels to provide sufficient amounts of the metabolites in a short period of time, thereby enabling a precise and quick analysis of the metabolites.

Moreover, they have also found that, despite that there are a considerable number of human cytochrome P450 molecular species, the human metabolic system for chemical compounds can be
20 reproduced in vitro when at least four human cytochrome P450 molecular species, i.e., human cytochrome P450 1A2, P450 2C9, P450 2E1 and P450 3A4, are combined.

Thus, the present invention provides a method for evaluation of the safety of a chemical compound, which comprises the steps of:

(a) reacting a chemical compound with recombinant yeast cells expressing, or in other words producing,
25 human cytochrome P450 molecular species P450 1A2, P450 2C9, P450 2E1 and P450 3A4 together with a yeast NADPH-P450 reductase, which may be in the form of a fused enzyme with each of said human cytochrome P450 molecular species, or with the cell free extracts of the yeast cells; and
(b) analyzing the resulting metabolite to determine the safety of the compound.

The present invention further provides a method for determination of the human metabolite of a
30 chemical compound, which comprises the steps of:

(a) reacting a chemical compound with recombinant yeast cells producing human cytochrome P450 molecular species P450 1A2, P450 2C9, P450 2E1 and P450 3A4 together with a yeast NADPH-P450 reductase, which may be in the form of a fused enzyme with each of said human cytochrome P450 molecular species, or with cell free extracts of the yeast cells; and
35 (b) identifying the resulting metabolite.

Figs. 1 to 4 show various primers for cloning human P450 genes.

Fig. 5 shows a synthetic linker for human P450 gene cloning.

Fig. 6 shows a method of constructing yeast expression plasmids for human P450 1A2.

Fig. 7 shows a method of constructing yeast expression plasmids for human P450 2C9.

40 Fig. 8 shows a method of constructing yeast expression plasmids for human P450 2E1.

Fig. 9 shows a method of constructing yeast expression plasmids for human P450 3A4.

Fig. 10 shows a method of constructing yeast expression plasmids for human P450 1A1.

Fig. 11 shows a method of constructing yeast expression plasmids for human P450 2A6.

Fig. 12 shows a method of constructing yeast expression plasmids for human P450 2B6.

45 Fig. 13 shows a method of constructing yeast expression plasmids for human P450 2C8.

Fig. 14 shows a method of constructing yeast expression plasmids for human P450 2C18.

Fig. 15 shows a method of constructing yeast expression plasmids for human P450 2C19.

Fig. 16 shows a method of constructing yeast expression plasmids for human P450 2D6.

Fig. 17 shows a method of constructing a yeast expression plasmid containing an artificial fused
50 enzyme gene.

Fig. 18 shows a method of constructing a yeast expression plasmid using a GAPDH promoter.

According to the present invention, it can be determined whether a test compound will be converted into a carcinogenic or mutagenic form through the metabolism in the human liver, and whether the test compound or its metabolite has mutagenicity.

55 Thus, the present invention provides a method for evaluation of safety of a chemical compound, and a method for determination of the human metabolite of a chemical compound.

Human Cytochrome P450 and Their Genes

The yeasts capable of expressing, or producing, said enzymes can be obtained by transforming them with expression plasmids containing genes encoding said enzymes with a conventional recombinant DNA method.

The human P450 molecular species to be used in the present invention include at least four human cytochrome P450 molecular species, i.e., human cytochrome P450 1A2, P450 2C9, P450 2E1 and P450 3A4. The genes encoding these essential human cytochrome P450 molecular species and yeast NADPH-P450 reductase are reported in Nucleic Acids Res., 14, 6773-6774, 1986; J. Biochem., 102, 1075-1082, 1987; J. Biol. Chem., 261, 16689-16697, 1986; DNA, 7, 79-86, 1988; and J. Biochem., 103, 1004-1010, 1988.

Although the kinds of P450 molecular species present in human liver vary among the race and individuals, the combination of said human P450 molecular species includes at least about 85% (molar ratio) of the total amount of the human P450 molecular species present in the human liver. Hence, the present method using the said combination of human P450 molecular species can accurately reproduce the human liver metabolism in vitro.

The combination of these P450 molecular species may optionally be varied, taking into account of the amounts of these P450 molecular species in the liver: the amount of P450 3A4 present in the human liver is about 35±10% of the total amount of the human P450 molecular species; P450 2C9 about 25±10%; P450 1A2 about 23±10%; and P450 2E1 about 17±10%.

In addition to the above-mentioned combination, human P450 molecular species P450 2A6, P450 2C19 and/or P450 2D6 (Biochemistry, 29, 1322-1329, 1990; Biochemistry, 30, 3247-3255, 1991; Am. J. Hum. Genet., 45, 889-904, 1989) may also be added. In this case, the combined human P450 molecular species covers at least 90% of the total amount of the human P450 molecular species present in the human liver.

The in vitro human metabolic system that reproduces accurately the human metabolism of a chemical compound, and can represent the differences among races and individuals can be obtained when these human P450 molecular species are properly combined, taking into account of the amounts of these species in the liver.

Furthermore, at least one human cytochrome P450 molecular species selected from the group of P450 1A1, P450 2B6, P450 2C8 and P450 2C18 (Science, 228, 80-83, 1985; Biochemistry, 28, 7340-7348, 1989; Nucleic Acids Res., 15, 10053-10054, 1987; Biochemistry, 30, 3247-3255, 1991) may be added to said human cytochrome P450 molecular species to reproduce in vitro the metabolism of the human liver more accurately.

The nucleotide sequences coding for the human P450 molecular species are disclosed in SEQ ID NOs: 1 to 38.

Cloning of Genes

The genes coding for the human cytochrome P450 molecular species are known and can be obtained by the conventional cloning methods.

For example, they may be obtained by:

- (i) preparing a mRNA fraction containing the mRNA of the gene coding for human cytochrome P450 molecular species;
- (ii) preparing a cDNA from the mRNA fraction using reverse transcriptase;
- (iii) preparing a cDNA library by inserting said cDNA into a phage vector or a plasmid vector; and
- (iv) cloning the gene coding for the human cytochrome P450 molecular species from the cDNA library obtained above or from a commercially available human liver-derived cDNA library using a DNA fragment having an identical sequence to some part of the desired gene or an antibody reactive to the protein produced by the gene.

The gene may also be obtained from the above-described cDNA library by the PCR method.

The gene coding for yeast NADPH-P450 reductase may be obtained by the same method as used for cloning of the genes coding for human P450 molecular species. More specifically, the gene may be obtained by such a known method as described in the Japanese Patent Laid-open Publication No. 62-19085.

Construction of Yeast Expression Plasmids

The yeasts capable of expressing said enzymes can be obtained by transforming them with expression plasmids containing genes encoding said enzymes with a conventional recombinant DNA method.

5 The yeast expression plasmid having a gene coding for human P450 molecular species and a gene coding for the yeast NADPH-P450 reductase can be constructed by using a conventional recombinant DNA method.

As to the promoter to be used for construction of the expression plasmids for the yeast of the present invention, there is no particular restriction so long as the promoter can be used in usual expression systems
10 for yeasts, and a promoter of a yeast alcohol dehydrogenase gene (hereinafter referred to as ADH promoter), glyceraldehyde-3-phosphate dehydrogenase promoter (hereinafter referred to as GAPDH promoter), and phosphoglycerate kinase (hereinafter referred to as PGK promoter) are preferably used in the present invention.

The ADH promoter can be prepared by a usual genetic engineering method, for example, from a yeast
15 expression vector pAAH5 possessing a yeast ADH1 promoter and terminator ("Methods in Enzymology" by Ammerer et al., vol.101, pp.192-201). The yeast ADH1 promoter is described in the U.S. Patent No. 299,733 to Washington Research Foundation and it requires patent license from the patentee in a case of using the same for an industrial or commercial purpose.

The yeast expression plasmid having both a gene coding for human P450 molecular species and a
20 gene coding for the yeast NADPH-P450 reductase can be constructed by, for example, inserting an NotI fragment prepared from yeast expression vector pAAH5N possessing the ADH promoter and terminator (Japanese Patent Laid-open Publication No. 2-211880) to an NotI site of plasmid pARRN possessing a gene coding for yeast NADPH-P450 reductase (Japanese Patent Laid-open Publication No. 2-211880) and then inserting cDNA coding for the human P450 molecular species to the HindIII site of the thus obtained
25 plasmid pAHRR. Moreover, a vector obtained by exchanging a Hind III site of pAAH5N with other restriction enzyme site may be used for the same purpose.

In the present invention a gene coding for an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase can also be used. The artificial fused enzyme can catalyze mono-oxygenation reaction and the efficiency of the electron transfer from NADPH is so
30 improved that the activity of the mono-oxygenation reaction is much enhanced. Accordingly, a great amount of metabolic products can be obtained in a shorter period of time, enabling accurate analysis.

The fused gene comprises a gene coding for the human cytochrome P450 molecule on the 5'-terminal and a gene coding for the yeast NADPH-P450 reductase on 3'-terminal.

The gene coding for such an artificial fused enzyme can be constructed by ligating a gene coding for a
35 human cytochrome P450 species and a gene coding for yeast NADPH-P450 reductase by a conventional recombinant DNA method, and the constructed gene is usually inserted to the Hind III site of the yeast expression vector pAAH5N having ADH promoter and ADH terminator described in the Japanese Patent Laid-open Publication No. 2-211880.

40 Transformation of Yeast

The yeast cells expressing the human P450 molecular species and yeast NADPH-P450 reductase or yeast cells expressing an artificial fused enzyme comprising human P450 molecular species and NADPH-P450 reductase can be obtained by introducing the thus constructed yeast expression plasmid into a yeast
45 by a known method such as a protoplast method or a method using alkaline metal salt (LiCl).

In the present invention, two or more expression plasmids may optionally be introduced into a single strain of yeast so that the yeast can express two or more molecular species simultaneously.

As the hosts, *Saccharomyces cerevisiae* is used in the method of the present invention, in particular, *Saccharomyces cerevisiae* AH22 (ATCC 38626) is preferably used.

50

Reaction of Test Compound

In the method of the present invention, a test compound is reacted with a mixture of at least said four human P450 molecular species, or separately with each of the said four human P450 molecular species if
55 the presence of the yeast NADPH-P450 reductase.

Alternatively, it may be first reacted with one or more of the essential human P450 molecular species, and then with a mixture of, or separately with the rest of them; each of the reactions is carried out in the presence of the yeast NADPH-P450 reductase.

The reaction is carried out by reacting a test compound with the yeast obtained by the transformation with an expression plasmid containing a gene encoding a human P450 molecular species and a gene encoding yeast NADPH-P450 reductase, or a fused gene encoding a fused enzyme of a human P450 molecular species and a yeast NADPH-P450 reductase, or with the cell free extracts of the yeast cells.

In the reaction of a test compound with the enzymes of the present invention, living yeast cells and their cell free extracts are usually used.

As the cell free extracts, subcellular fraction of cells containing microsomal fractions, or fractions containing both microsome and cytoplasm is used. The cell free extracts or fractions can be prepared, for example, by a known method (DNA, Vol.4, No. 3, pp.203-210 (1985)).

However, the present invention can be preferably carried out with the cell free extracts, especially with microsomal fractions of the cells. But, when biological analytic method is used to determination of the mutagenicity or carcinogenicity, fractions containing microsome and cytoplasm are preferably used.

The reaction can be conducted by adding a test compound to a culture solution or a buffer solution of yeast cells or cell free extracts, and the resultant solution is usually incubated at a temperature, for example, at about 10 °C to 40 °C, for about 0.1 to 48 hours.

The amounts of the yeast cells or cell free extracts and the compound vary depending on the conditions such as reaction temperature, reaction time and the kind of the test compound to be used.

For instance, the amount of the yeast cells to be used in the solution is preferably from about 10^5 to about 10^{10} per 1 ml of the solution, preferably, from about 10^7 to about 10^8 per 1 ml of the solution. When cell free extracts are used, from about 10^{10} to about 10^{15} of P450 molecules per 1 ml of the solution, preferably from about 10^{12} to about 10^{13} of P450 molecules per 1 ml of the solution is usually used.

The amount of the compound to be added is preferably within a range of from about 0.01 μ mol to about 1 μ mol per 1 ml of the solution.

The above ranges may be optionally varied, if necessary.

Determination of Metabolites

The metabolites present in the reaction solution can then be subjected to elucidation of the chemical structures and the measurement of their amounts. The analysis of the chemical structure can be conducted by known methods ("Guide to Apparatus Analysis (2)", edited by Jiro Shiokawa et al., (revised edition) first print, issued from Kagaku Dojin (1985); "Spectral Identification for Organic Compound" by R.M. Silverstein, fourth edition, third print, issued from Tokyo Kagaku Dojin (1984)).

From the results of the analysis of the metabolites, it can be determined whether the tested compound will be detoxicated or metabolized into a carcinogen in the human liver when administered.

Determination of Toxic Effects of Metabolites

The toxic effects, in particular mutagenicity, of the resulting metabolites can be determined by a conventional biological analytic method such as the Ames Test. For example, the metabolites present in the reaction solution are allowed to react with mutant bacteria such as histidine requiring Salmonella strain (Salmonella typhimurium (his-)), or tryptophan requiring Escherichia coli (Escherichia coli (trp-)), and then determine whether the metabolites cause the back mutation of the bacteria whether the colonies of revertant which is not requiring the amino acid (His+ or Trp+) are formed, and, if formed, how many colonies. In place of the bacteria, mammalian cells such as MCL-5 cells, which are sensitive to cell toxicity of a chemical compound (U.S. Patent No. 4,532,204), can be used.

In this method, the compounds that cause the back mutation will be judged to be mutagenicity test-positive.

It is also possible to simultaneously proceed the step (a) of reacting the test compound with the yeast cells or the cell free extracts, and the step (b) of analyzing the metabolites present in the reaction solution.

The mutagenicity of arylamine derivatives, which are known to be metabolized by the liver into a mutagens, can be examined by the biological analytic method. For example, the mutagenicity of 2-aminoanthrathene can be detected at the concentration of about 0.1 μ g of 2-aminoanthrathene when 20 pmol of P450 1A2, which is active specifically to 2-aminoanthrathene, is used (Table 1).

In the present invention, a metabolic probe for a human P450 molecular species can be obtained.

If a certain chemical compound is converted by a particular human P450 molecular species into a specific metabolite, the amount of such a human P450 molecular species can be determined by detecting such a metabolite in excretions such as blood or urine of a living body who has been administered the compound, and such a compound is called a metabolic probe.

In the present invention, such a metabolic probe can be obtained by screening the metabolites obtained by reacting chemical compounds with the yeasts of the present invention.

Examples

The present invention will be further illustrated by the following examples, which are not to be construed to limit the scope thereof.

Preparation of cDNA coding for human P450 molecular species

cDNA coding for human P450 molecular species were obtained from commercially available human liver cDNA library (Clontech Co.) by the PCR method using primers for cloning human P450 genes as shown in Figs. 1 to 4, and a method using a synthetic linker for human P450 gene cloning as shown in Fig. 5. Thus obtained nucleotide sequences for the cDNA and the deduced amino acid sequences are shown in the sequence listing.

Relationship between SEQ ID NOs and human P450 molecular species are as follows:

1. The essential human cytochrome P450 molecular species for the present invention.

(1) SEQ ID NO: 1	1A2
(2) SEQ ID NO: 3	2C9
(3) SEQ ID NO: 5	2E1
(4) SEQ ID NO: 7	3A4

2. Auxiliary Human cytochrome P450 molecular species

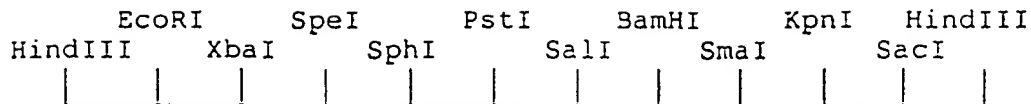
(1) SEQ ID NOs: 9, 11 and 13	1A1
(2) SEQ ID NOs: 15 and 17	2A6
(3) SEQ ID NO: 19	2B6
(4) SEQ ID NOs: 21, 23 and 25	2C8
(5) SEQ ID NO: 27	2C18
(6) SEQ ID NO: 29	2C19
(7) SEQ ID NOs: 31, 33, 35 and 37	2D6

Construction of yeast expression plasmids: p1A2 and p1A2R

Fig. 6 shows a method of constructing yeast expression plasmids for human P450 1A2. The protein coding region of P450 1A2 gene of about 1.5 kb excluding about 40 bp at the 5'-terminal was amplified by the PCR method using the primers shown in Fig. 1. The resultant fragment of about 1.5 kb was cleaved with SacI and sub-cloned to a pUC118 vector. About 40 bp at the 5'-terminal was chemically synthesized as the linkers shown in Fig. 5 and sub-cloned between the HindIII and SacI sites of the pUC118 vector. The plasmid having the 1.5 kb fragment was digested by HindIII, blunted, and then ligated with an EcoRI linker. The EcoRI-SacI fragment was prepared from the resulting plasmid and ligated into the plasmid containing the 5'-terminal 40 bp. Then, it was treated with EcoRI and blunted. A HindIII linker was inserted into the blunted fragment. The obtained fragment then cleaved with HindIII was inserted into pAAH5N and pAHRP to construct a yeast expression plasmid p1A2 for human P450 1A2, and a yeast expression plasmid p1A2R for simultaneous expression of human P450 1A2 and yeast NADPH-P450 reductase.

Construction of yeast expression plasmids: p2C9 and p2C9R

Fig. 7 shows a method of constructing yeast expression plasmids for human P450 2C9. The protein coding region of 450 2C9 gene was divided into two fragments of about 0.9 kb and about 0.6 kb, and the fragments were amplified by the PCR method using the primers shown in Fig. 1. The resultant fragment of about 0.9 kb was cleaved with PstI and sub-cloned to a pUC B vector, which was prepared by exchanging the cloning site located between the two Hind III sites, one of which was obtained by converting the EcoRI site of pUC19, with the following cloning sites:



5

The fragment of about 0.6 kb was incorporated between the XbaI and PstI sites of the plasmid having the 0.9 kb fragment to ligate the two segments. The KpnI site of the plasmid was blunted. An XbaI linker was inserted to the blunted plasmid. The XbaI fragment containing the coding region was cut out from the resultant fragment. A modified pUC vector, pUCAN, was constructed by replacing the EcoRI and HindIII sites with NotI sites, followed by insertion of the NotI fragment prepared from pAAH5N between the two NotI sites. The HindIII site of pUCAN vector having the ADH promoter and terminator regions in the pUC vector was blunted and inserted into pUCANX introduced with the XbaI linker. The obtained plasmid was cleaved with NotI and inserted into pAAH5N and pAHRR treated in a similar manner with NotI to construct a yeast expression plasmid p2C9 for human P450 2C9, and a yeast expression plasmid p2C9R for simultaneous expression of human P450 2C9 and yeast NADPH-P450 reductase.

Construction of yeast expression plasmids: p2E1 and p2E1R

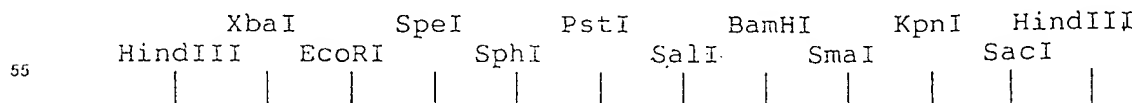
Fig. 8 shows a method of constructing yeast expression plasmids for human P450 2E1. The protein coding region of P450 2E1 gene was divided into two fragments of about 0.5 kb and about 1.0 kb, both of which were amplified by the PCR method using the primers shown in Fig. 1. The resultant fragment of about 0.5 kb was cleaved with EcoRI and BamHI and sub-cloned to a pUC118 vector. Then the fragment of about 1.0 kb was incorporated between the BamHI and SphI sites to ligate the two fragments. This was cleaved with EcoRI, and SphI, and inserted into pUC B first and then cut out with HindIII. The resultant fragment was inserted into pAAH5N and pAHRR vectors to construct a yeast expression plasmid p2E1 for human P450 2E1, and a yeast expression plasmid p2E1R for simultaneous expression of human P450 2E1 and yeast NADPH-P450 reductase.

Construction of yeast expression plasmids: p3A4 and p3A4R

Fig. 9 shows a method of constructing yeast expression plasmids for human P450 3A4. The protein coding region of P450 3A4 gene was divided into two fragments of about 0.6 kb and about 0.9 kb, both of which were amplified by the PCR method using the primers shown in Fig. 2. The resultant fragment of about 0.6 kb was cleaved with SacI and sub-cloned to a pUC118 vector. Subsequently, it was cleaved with EcoRI and blunted. An XbaI linker was ligated to the blunted fragment. The fragment of 0.9 kb was cleaved with XbaI and SacI, and incorporated to the resultant fragment above, thus the two fragments were ligated. After cleaving the plasmid with SphI, it was blunted. An XbaI linker was ligated to the blunted fragment, from which the XbaI segment was cut out and inserted to an XbaI site of pUCANX. This was cut out with NotI and inserted into pAAH5N and pAHRR treated in a similar manner with NotI. Thus a yeast expression plasmid p3A4 for human P450 3A4, and a yeast expression plasmid p3A4R for simultaneous expression of human P450 3A4 and yeast NADPH-P450 reductase were constructed.

Construction of yeast expression plasmids: p1A1 and p1A1R

Fig. 10 shows a method of constructing yeast expression plasmids for human P450 1A1. The coding region for P450 1A1 protein was divided into two fragments of about 1.0 kb and about 0.5 kb and the resultant fragments were amplified by the PCR method using the primers shown in Fig. 2. Thus obtained fragment of about 1.0 kb was cleaved with XbaI and SacI and sub-cloned to a PUC19 vector, which was prepared by exchanging the cloning site located between the two HindIII sites, one of which was obtained by converting the EcoRI site of pUC19, with the following cloning sites:



55

The amplified fragment of about 0.5 kb was sub-cloned into the HincII site of a pUC 19 vector and the resultant plasmid was then cleaved with SacI. The cleaved fragment was ligated with the plasmid having the 1.0 kb fragment. After cutting out the coding region from the thus obtained 1A1 gene with HindIII, the fragment was inserted to the HindIII site of the yeast expression vector pAAH5N having ADH promoter and terminator regions, and to the same site of vector pAHRR for simultaneous expression of P450 and yeast NADPH-P450 reductase of which gene is located upstream of the P450 gene. Thus yeast expression plasmid p1A1 for human P450 1A1 and yeast expression plasmid p1A1R for simultaneous expression of human P450 1A1 and yeast NADPH-P450 reductase were constructed.

In addition two kinds of human P450 1A1 gene fragments which were different only in a small portion of the nucleotide sequence were obtained in a similar manner and used to construct two kinds of yeast expression plasmid for human P450 1A1, p1A1 Variant 1 and p1A1 Variant 2, and two kinds of plasmids for simultaneous expression of human P450 1A1 and yeast NADPH-P450 reductase, p1A1R Variant 1 and p1A1R Variant 2.

Construction of yeast expression plasmids: p2A6 and p2A6R

Fig. 11 shows a method of constructing yeast expression plasmids for human P450 2A6. A protein coding region of P450 2A6 gene was divided into two fragments of about 0.6 kb and about 0.9 kb, both of which were amplified by the PCR method using the primers shown in Fig. 2 to yield two kinds of human P450 2A6 gene fragments which were different only in a small portion of the nucleotide sequence. The resultant fragment of about 0.6 kb was cleaved with XbaI and HincII, and sub-cloned to a pUC A vector. Then the fragment of 0.9 kb was incorporated between the HincII and KpnI sites to ligate the two fragments. The obtained fragment was cleaved with HindIII and inserted into pAAH5N and pAHRR to construct two kinds of yeast expression plasmid for human P450 2A6, p2A6 and p2A6 Variant 1, and two kinds of yeast expression plasmid for simultaneous expression of human P450 2A6 and yeast NADPH-P450 reductase, p2A6R and p2A6R Variant 1.

Construction of yeast expression plasmids: p2B6 and p2B6R

Fig. 12 shows a method of constructing yeast expression plasmids for human P450 2B6. The entire protein coding region of P450 2B6 gene was amplified by the PCR method using the primers shown in Fig. 3. The resultant fragment was cleaved with XbaI and BamHI and sub-cloned to pUC A. The resulting plasmid was partially digested with HindIII, and inserted into pAAH5N and pAHRR vectors to construct a yeast expression plasmid p2B6 for human P450 2B6, and a yeast expression plasmid p2B6R for simultaneous expression of human P450 2B6 and yeast NADPH-P450 reductase.

Construction of yeast expression plasmids: p2C8 and p2C8R

Fig. 13 shows a method of constructing yeast expressed plasmids for human P450 2C8. The entire protein coding region of the P450 2C8 gene was amplified by the PCR method using the primers shown in Fig. 3 to yield three kinds of P450 2C8 genes which were different only in a small portion of the nucleotide sequence. The resultant fragments were partially digested with XbaI, and sub-cloned to pUC A. The fragment was cleaved with HindIII and inserted into pAAH5N and pAHRR vectors to construct three kinds of yeast expression plasmids p2C8, p2C8 Variant 1 and p2C8 Variant 2 for human P450 2C8, and three kinds of yeast expression plasmids, p2C8R, p2C8R Variant 1 and p2C8R Variant 2 for simultaneous expression of human P450 2C8 and yeast NADPH-P450 reductase.

Construction of yeast expression plasmids: p2C18 and p2C18R

Fig. 14 shows a method of constructing yeast expression plasmids for human P450 2C18. The protein coding region of P450 2C18 gene was divided into two segments of about 1.4 kb and about 0.9 kb, then the both fragments were amplified by the PCR method using the primers shown in Fig. 3. The amplified fragment of about 1.4 kb was cleaved with PstI and sub-cloned to a pUC A vector. The fragment of about 0.9 kb was incorporated between the XbaI and PstI sites to ligate the two fragments. After cleaving the plasmid with SmaI, an XbaI linker was introduced. Then an XbaI fragment was prepared and inserted into the XbaI site of pUCANX. It was cleaved with NotI and inserted into pAAH5N and pAHRR treated in a similar manner with NotI to construct a yeast expression plasmid p2C18 for human P450 2C18, and a yeast expression plasmid p2C18R for simultaneous expression of human P450 2C18 yeast and NADPH-P450

reductase.

Construction of yeast expression plasmids: p2C19 and p2C19R

5 Fig. 15 shows a method of constructing yeast expression plasmids for human P450 2C19. Fragments a, b and c for the protein coding region of P450 2C19 gene were amplified by the PCR method using the primers No. 1, No. 2, No. 3 and No. 4, No.5 and No. 6, and No.5 and No.7 defined by SEQ ID NOs: 39-45, respectively.

10 Fragments e and f for the protein coding region of human cytochrome P450 2C19 were also amplified against human cytochrome P450 2C9 gene by the PCR method using the primers No. 8 to 21 having nucleotide sequences with some mutations shown by SEQ ID NOs: 46 to 59. A fragment d for the linker Nos. 1 and 2 having nucleotide sequences shown by SEQ ID NOs: 60 and 61 was obtained by directly synthesizing the DNA to cover the rest of the protein coding region of the human P450 2C19 gene. Thus the fragments covering the whole protein coding region of the human cytochrome P450 2C19 were
15 obtained.

After the fragments a and b were treated with XhoI and BamHI, and with BamHI and PstI, both fragments were simultaneously inserted between the XhoI and PstI sites of the Blue Script(+). The fragment e was treated with XbaI and XhoI and inserted to the XbaI and XhoI sites of the plasmid having the fragments a and b to give a plasmid having the fragments a, b and e.

20 After the fragment c was treated with PstI and KpnI, the resulting fragment was simultaneously inserted with the linker fragment d between the PstI and EcoRI sites of the Blue Script(+). The resultant plasmid was cut with PstI and EcoRI to give a fragment containing the fragments c and d. Then this fragment was simultaneously inserted between the fragment f treated with EcoRI to the PstI and HincII sites of the aforementioned plasmid containing the fragment a, b and e. Thus a plasmid having the whole coding region
25 of the human cytochrome P450 2C19 gene was constructed. The constructed plasmid was cut with HindIII and the resultant fragment was inserted to pAAH5N and pAHRR both of which were treated with HindIII to give a yeast expression plasmid p2C19 for expressing the human P450 2C19 and a yeast expression plasmid p2C19R for simultaneous expression of the human P450 2C19 and yeast NADPH-P450 reductase.

30 SEQ ID NOs and primer Nos. are as follows:

35	SEQ ID No: 39	Primer No. 1
	SEQ ID NO: 40	Primer No. 2
	SEQ ID NO: 41	Primer No. 5
	SEQ ID NO: 42	Primer No. 4
	SEQ ID NO: 43	Primer No. 5
	SEQ ID NO: 44	Primer No. 6
40	SEQ ID NO: 45	Primer No. 7
	SEQ ID NO: 46	Primer No. 8
	SEQ ID NO: 47	Primer No. 9
	SEQ ID NO: 48	Primer No. 10
	SEQ ID NO: 49	Primer No. 11
	SEQ ID NO: 50	Primer No. 12
45	SEQ ID NO: 51	Primer No. 13
	SEQ ID NO: 52	Primer No. 14
	SEQ ID NO: 53	Primer No. 15
	SEQ ID NO: 54	Primer No. 16
	SEQ ID NO: 55	Primer No. 17
50	SEQ ID NO: 56	Primer No. 18
	SEQ ID NO: 57	Primer No. 19
	SEQ ID NO: 58	Primer No. 20
	SEQ ID NO: 59	Primer No. 21
	SEQ ID NO: 60	Linker No. 1
55	SEQ ID NO: 61	Linker No. 2

Construction of yeast expression plasmids: p2D6 and p2D6R

Fig. 16 shows a method of constructing yeast expression plasmids for human P450 2D6. The protein coding region of 1.3 kb excluding about 200 bp at the 5'-terminal of P450 2D6 gene was divided into two fragments of about 0.4 kb and about 0.9 kb, and the both fragments were amplified by the PCR method. The resultant fragment of about 0.9 kb was cleaved with KpnI and sub-cloned to pUC A. For the 200 bp on the 5'-terminal, three synthetic linkers shown in Fig. 5 were used and two linkers on the 5'-terminal were incorporated into XbaI and PstI sites of a Blue Script(+) vector and then other linkers were incorporated into SmaI and PstI sites. Then fragment of about 0.4 kb obtained by the PCR method was incorporated into the PstI and HincII sites of the plasmid and then cleaved with NspV and XbaI. The resultant fragment was inserted into the plasmid containing the 0.9 kb fragment to ligate the coding region. This was cleaved with HindIII and inserted into pAAH5N and pAHRH vectors to construct a yeast expression plasmid p2D6 for human P450 2D6, and a yeast expression plasmid p2D6R for simultaneous expression of human P450 2D6 and yeast NADPH-P450 reductase.

Then three kinds of human P450 2D6 gene fragments which were different only in a small portion of the nucleotide sequence were obtained in a similar manner as described above and used to construct two kinds of yeast expression plasmids for human P450 2D6, p2D6 Variant 1, p2D6 Variant 2 and p2D6 Variant 3, and three kinds of yeast expression plasmid 2D6R for simultaneous expression of human P450 2D6 yeast and NADPH-P450 reductase, p2D6R Variant 1, p2D6R Variant 2 and p2D6R Variant 3.

Construction of yeast expression plasmid containing artificial fused enzyme gene

An expression plasmid was constructed in accordance with Fig. 17. The XbaI-XhoI fragment was amplified with plasmid p3A4 by using the primers shown in Fig. 4. On the other hand, the XhoI-HindIII fragment of about 2.1 kb was obtained from the plasmid pBFCRI (Japanese Patent Application No. 4-209226) and inserted between the XhoI and HindIII sites of a commercial vector Blue Script(+), followed by digestion with restriction enzymes XhoI and XbaI. These two fragments were simultaneously inserted to the XbaI site of the vector pUCAN, which was then digested with NotI to give a fragment of about 5.6 kb. The desired yeast expression plasmid pF3A4 was obtained by ligating the fragment with the NotI fragment of about 10.5 kb obtained from vector pAAH5N (Japanese Patent Laid-open Publication No. 2-211880). The artificial fused enzyme consists of 1156 amino acid residues of which sequence structure comprising, successively, from the N-terminal end, an entire amino acid sequence (503 residues) of human liver cytochrome P450 3A4, a linker-derived sequence (Ala-Arg-Ala), and a sequence of from the 42nd residue to C-terminal of yeast NADPH-cytochrome P450 reductase.

Preparation of transformed yeast cell

Saccharomyces cerevisiae AH 22 was inoculated to 1.0 ml of YPD culture medium (1% yeast extract, 2% polypeptone, 2% glucose). After shaken at 30 °C for 18 hours, the yeast cells were collected by centrifugation (5000 x g, 10 min). The resultant cells were suspended in 10 ml of 0.2 M LiCl solution and then centrifuged again (5000 x g, 10 min) to obtain pellets. Then 20 µl of 1 M LiCl solution, 30 µl of 70% polyethylene glycol 4000 and each 10 µl solution containing about 1.0 µg of various kinds of yeast expression plasmids for the human P450 molecular species and yeast NADPH-reductase constructed as above were added to the resultant pellets. After sufficiently mixing them, they were incubated at 30 °C for one hour and further stirred after the addition of 140 µl of sterilized water. The solution was plated on SD synthetic culture medium (2.0% glucose, 0.67% nitrogen base w/o amino acids, manufactured by Difco Co., 20 µg/ml of histidine, 2.0% agar) and incubated at 30 °C for three days. Then transformed yeast cells possessing the yeast expression plasmid described above were selected. In this way, various kinds of yeast cells expressing the human P450 molecular species were prepared.

Quantitative measurement of human P450 expressed in yeast

Each 200 ml of culture broth of each kind of yeast cells expressing human P450 molecular species and yeast NADPH-reductase simultaneously or expressing an artificial fused enzyme comprising human P450 molecular species and yeast NADPH-reductase prepared as above (SD synthetic culture medium, cell concentration: about 1.5×10^7 cells/ml) was used to collect the cells. The collected cells were then suspended in 10 ml of 100 mM potassium phosphate buffer solution (pH 7.0) and centrifuged (5000 x g, 10 min) to obtain pellets. Thus obtained pellets were resuspended in 2.0 ml of 100 mM potassium phosphate

buffer solution (pH 7.0) and 1 ml of each of the solutions were poured into two cuvettes. After bubbling carbon monoxide to a sample cuvette, 5 to 10 mg of dithionite was added to both of the cuvettes, and stirred and then difference spectrum at 400-500 nm was measured to calculate the concentration of P450 present in the yeast. The amount of each kind of human P450 species or an artificial fused enzyme in each kind of transformed yeast cells was at a level from about 10^5 to about 10^6 molecules/cell.

Preparation of yeast S-9 Mix fraction, cytoplasmic fraction and microsomal fraction

First, 3.8 liter of each kind of culture broth (SD synthetic culture medium, cell concentration: about 1.0×10^8 cells/ml) of yeast cells expressing human P450 molecular species and yeast NADPH-reductase simultaneously or an artificial fused enzyme comprising human P450 molecular species and yeast NADPH-reductase prepared as above was collected and the resultant cells were suspended in 400 ml of a buffer solution A (10 mM Tris-HCl (pH 7.5), 2 M sorbitol, 0.1 mM DTT, 0.2 mM EDTA), to which 160 mg of Zymolyase 100,000 (Zymolyase 100T) was added, and the obtained solution was incubated at 30°C for 60 min. Spheroplast obtained by centrifugation ($5000 \times g$, 10 min) was suspended in 100 ml of the buffer solution A and then centrifuged ($5000 \times g$, 10 min). Washing the spheroplast by repeating the same centrifugal operation once again, the spheroplast was finally suspended in 200 ml of a buffer solution (10 mM Tris-HCl (pH 7.5), 0.65 M sorbitol, 0.1 mM DTT), which was then subjected to ultrasonic pulverization (50 W, for 5 min). The cell free extracts were centrifuged ($9000 \times g$, 20 min) and supernatants were recovered to obtain a yeast S-9 Mix fraction. Further, the fraction was centrifuged ($125,000 \times g$, 70 min) to collect precipitates which were suspended again into 10 ml of 0.1 M potassium phosphate buffer solution (pH 7.4) to obtain a microsomal fraction. On the other hand, a cytoplasmic fraction was obtained by recovering the supernatants.

Construction of yeast expression plasmid using GAPDH promoter and its expression in yeast

Fig. 18 shows a method of constructing a yeast expression plasmid using a GAPDH promoter. A HindIII fragment (about 3.0 kb) obtained from pARRN (described in the Japanese Patent Laid-open Publication No. 2-211880) was inserted into a HindIII site of plasmid pUN, which was obtained by cleaving pUC19 with EcoRI, blunt-ending and ligation with an NotI linker to give pUR. On the other hand, after blunting an XhoI site of plasmid pAAH5 and inserting an XbaI linker, it was cleaved with restriction enzymes XbaI and Sall and the resultant fragment (about 2.2 kb) was inserted to XbaI and Sall sites of pUC19. The three fragments, namely, a fragment (about 2.2 kb) obtained by cleaving the resultant plasmid with XbaI and PstI, the XbaI-PstI fragment (about 1.3 kb) cut out from 2 μm DNA of *Saccharomyces cerevisiae* AH22, and a fragment obtained by cleaving pUR with PstI were ligated to give a plasmid pURL. Further, the pURL was cleaved with HindIII, blunted and ligated to remove the HindIII site. Then, an NotI fragment (about 1.6 kb) containing GAPDH promoter and terminator (obtained by the method as described in Agric. Biol. Chem., 51, 1641-1647 (1987) and J. Biol. Chem., 267, 16497-16502 (1992)) was ligated to the NotI site of pURL to give a plasmid pURLG. Human P450 2D6 cDNA obtained by the method used for the construction of p2D6 was inserted to a HindIII site of pURLG to obtain a yeast expression plasmid pG2D6R for simultaneous expression of human P450 2D6 and yeast NADPH-P450 reductase. When the plasmid was introduced by the method used in the preparation of transformed yeast cells as above to *Saccharomyces cerevisiae* AH22, production of human P450 2D6 was observed.

Metabolism of 7-ethoxycoumarin using transformed yeast cells

7-Ethoxycoumarin was added to each 2 ml of the culture media of the transformed yeast cells expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase (SD synthetic culture medium, cell concentration: about 2.0×10^7 cells/ml) so that the final concentration of 7-ethoxycoumarin was 0.5 mM. After incubation at 30°C for 2 or 5 hours, supernatants were obtained by centrifugation ($5000 \times g$, 10 min). To the supernatants 62.5 μl of 15% TCA (trichloroacetic acid) and 2 ml of chloroform were added and, after well stirring, a chloroform layer was recovered by centrifugation ($5000 \times g$, 10 min), to which 4 ml of 0.01 N NaOH containing 0.1 M NaCl was added and stirred sufficiently and then centrifuged ($5000 \times g$, 10 min). After recovering the supernatants, fluorescence was measured for the supernatant fraction (ex. 366 nm, em 452 nm) to quantitatively measure the reaction product 7-hydroxycoumarin. As a result, O-deethylation activity for 7-ethoxycoumarin can be observed for all of 11 kinds of the yeast cells expressing the human P450 molecular species. P450 1A1 and P450 2B6

showed strong activity; and P450 1A2, P450 2E1, P450 2A6 and P450 2D6 showed good activity, while P450 2C8, P450 2C9, P450 3A4, P450 2C18 and P450 2C19 showed moderate activity.

Metabolism of tolbutamide using transformed yeast cells

5

In the same manner as above, tolbutamide was added to each of the culture solutions of the transformed yeast cells expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase so that the concentration of the compound was 1.0 mM. After incubation at 30 °C for 15 hours, the culture supernatant was then obtained by centrifugation (5000 x g, 10 min). To the supernatant, 2 ml of dichloromethane was added. After sufficient stirring, the dichloromethane layer was recovered by centrifugation (5000 x g, 10 min), and the solvent was evaporated under reduced pressure. The resultant residue was dissolved in 100 µl of acetonitrile, and the solution was analyzed by HPLC under the following conditions. As a result, hydroxylated tolbutamide was detected in the solution of yeast cells expressing human P450 2C8, P450 2C9, P450 2C18 and P450 2C19. The human P450 2C9 showed high activity and 2C19 showed good activity. On the other hand, hydroxylated tolbutamide was not detected in the solution of yeast cells expressing other human P450 than described above.

Conditions for HPLC

Column: µBondapak C18 (manufactured by Waters Co.)
 20 Carrier: 10-70% acetonitrile-distilled water (linear concentration gradient for 20 min)
 Temperature: 50 °C
 Detection: UV 230 nm
 Injection amount: 50 µl

25 Metabolism of testosterone using transformed yeast cells

In the same manner as above, testosterone was added to each of the culture solutions of the transformed yeast cells expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase so that the concentration of the compound was 0.05 mM. After incubation at 30 °C for 15 hours, the supernatant was obtained by centrifugation (5000 x g, 10 min). Then 2 ml of dichloromethane was added. After sufficient stirring, the solution was centrifuged again (5000 x g, 10 min). The dichloromethane layer was recovered from the separated layer and the solvent was evaporated under reduced pressure. The resultant residue was dissolved in 100 µl of acetonitrile, and the solution was analyzed by HPLC under the following conditions. As a result, hydroxylated testosterone was detected for yeast cells expressing human P450 1A1, P450 2C8 and P450 3A4. On the other hand, hydroxylated testosterone was not detected for yeast cells expressing other human P450 than described above.

Conditions for HPLC

Column: µBondapak C18 (manufactured by Waters Co.)
 40 Carrier: 20-70% acetonitrile-distilled water (linear concentration gradient for 25 min)
 Temperature: 50 °C
 Detection: UV 254 nm
 Injection amount: 50 µl

45 Metabolism of chlorzoxazone using transformed yeast cells and microsomal fractions thereof

Chlorzoxazone was added to each of the culture solutions of the transformed yeast cells expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase as above so that the concentration of the compound was 0.5 mM. After incubation at 30 °C for 15 hours, the supernatant was obtained by centrifugation (5000 x g, 10 min). Then 2 ml of dichloromethane was added to the supernatant and vigorously stirred and centrifuged (5000 x g, 10 min). The dichloromethane layer was recovered from the separated layer, then evaporated under reduced pressure. The obtained residue was dissolved in 100 µl of acetonitrile, and the solution was analyzed by HPLC under the following conditions.

55 NADPH and chlorzoxazone were added to a microsomal fraction of yeasts expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase prepared as above so that the concentrations of NADPH and chlorzoxazone were 0.5 mM and 250 µM. Then the

solutions were incubated at 37 °C for 10 min. After that, trichloroacetic acid was added to the solutions so that the concentration of the trichloroacetic acid was about 10% (v/v). Then 2 ml of dichloromethane was added to the solution, and the solution was stirred vigorously and centrifuged (15,000 x g, 5 min). The dichloromethane layer was recovered, and the solvent was removed under reduced pressure. The obtained residue was dissolved in 100 µl of acetonitrile and the solution was subjected to analysis by HPLC under the same conditions as above.

All of the yeast cells expressing eleven human P450 molecular species gave hydroxylated chlorzoxazone. P450 2E1 showed high activity, and P450 1A1, P450 1A2, P450 2A6, P450 2D6 showed good activity, while P450 2C8, 2C9, 2B6, 2C18, 2C19 and 3A4 showed moderate activity.

Ames test using yeast S-9 Mix fraction and microsomal fraction

The Ames test method was in accordance with the customary method described, for example, in Mutat. Res., (1975) 31, 347. 2-Aminoanthracene which is an arylamine type compound was used as a specimen compound. (1) Rat S-9 Mix supernatant fraction (obtained by homogenizing liver and then subjected to centrifugation (9000 x g, 10 min), manufactured by Kikkoman) containing each kind of rat P450 molecular species at the concentration of 1200 pmol per 1 sample and (2) Yeast S-9 Mix fraction obtained from each kind of yeast cells expressing human P450 or a microsomal fraction prepared from the yeast S-9 Mix fraction were used as a metabolic activation source in the Ames test. As a result, more than 1000 revertant colonies were detected for the compound at 1 µg/plate (90 mm dia.) only in the case of using the yeast S-9 Mix fraction obtained from the yeast cells expressing human P450 1A2 (*Saccharomyces cerevisiae* AH22/p1A2R) and yeast cells expressing human P450 2E1 (*Saccharomyces cerevisiae* AH22/p2E1R) and a microsomal fraction prepared from the yeast S-9 Mix fraction, while the amounts of the human P450 molecules of these fractions were only one five hundredth and one thirtieth of the human P450 molecules present in the Rat S-9 mixture.

The human cytochrome P450 1A2 showed high activity, and human P450 2E1 showed only moderate activity. But the revertant colonies were not found for the human cytochrome P450 3A4, 2C8 and 2A6.

Metabolism of acetanilide using transformed yeast cells

Acetanilide was added to each of the culture solutions of the transformed yeast cells expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase, so that the concentration of the compound was 5 mM, and the solution was incubated at 30 °C for 15 hours. Then the solution was centrifuged (5000 x g, 10 min) to give a supernatant. The obtained supernatant solution was subjected to the HPLC analysis under the following conditions. The hydroxylated acetanilide was found for all of the tested eleven human P450 molecular species.

Among them, P450 1A2 and 2D6 showed high activity and P450 1A1, 2A6, 2B6, 2C8, 2C9, 2C18, 2C19 and 2E1 showed good activity, while 3A4 showed moderate activity.

Conditions for HPLC

Column:	µBondapak C18 (manufactured by Waters Co.)
Carrier:	Methanol:water:acetic acid = 15:84:1
Temperature:	30 °C
Detection:	UV 254 nm
Injection amount:	50 µl

Metabolism of coumarin using transformed yeast cells

Coumarin was added to 6 ml of each of the culture solutions (SDS synthetic culture medium, cell concentration of about 2.0×10^7 cells/ml) of the transformed yeast cells expressing (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase prepared as above, so that the concentration of the compound was 5 mM, and the solution was incubated at 30 °C for 2 or 5 hours. Then the solution was centrifuged (5000 x g, 10 min) to give a supernatant. 62.5 µl of 15% trichloroacetic acid and 2 ml of chloroform were added to the obtained supernatant solution, and the resultant solution was stirred well. The chloroform layer was recovered from the separated layer. Then 4 ml of sodium hydroxide solution containing 0.1 M NaCl was added to the solution and centrifuged again (5000 x g, 10 min). The supernatant fraction was recovered and subjected to fluorescence analysis (ex. 366 nm, em. 452 nm) to

measure the 7-hydroxycoumarin formed. The hydroxylation activity was specifically found only for the yeast cells expressing the human P450 2A6, while other yeast cells showed no activity.

Metabolism of debrisoquine using the microsomal fraction of transformed yeast whole cells

5 NADPH and [14 C]debrisoquine were added to each microsomal fraction solution of (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase prepared as above, so that the concentration of the compound was 100 μ M and that of NADPH is 6 mM, and the
10 solution was incubated at 30 °C for 30 minutes. Then perchlorate was added to the solution, so that the final concentration of the perchlorate was 10% (v/v). The solution was sufficiently stirred and centrifuged (15,000 x g, 15 min) to give the supernatant. The obtained supernatant was subjected to HPLC analysis according to the following conditions.

15 Microsomal fractions of yeasts expressing P450 1A1 and 2D6 showed good activity for the hydroxylation of the debrisoquine, while those of yeast cells expressing other human P450 molecular species showed no activity.

Conditions for HPLC

Column: COSMOSIL 5C18 (manufactured by Nakarai Tesq Co.)
Carrier: A:(acetonitrile)/B:(20mM Sodium Perchlorate, pH = 2.5)

20

Time (minute)	A/B
0-15	9/91
15-30	9/91-25/75 (linear gradient)
30-32	100/0
32-42	9/91

25

Temperature: room temperature

Detector: RI 14 C

30

Injection amount: 100 μ l

Metabolism of S-mephenytoin using the microsomal fraction of transformed yeast cells

35 NADPH and [14 C]S-mephenytoin were added to each microsomal fraction solution of (i) human cytochrome P450 molecular species and yeast NADPH-P450 reductase; or (ii) an artificial fused enzyme comprising human cytochrome P450 molecular species and yeast NADPH-P450 reductase prepared as above, so that the concentration of the compound was 25 μ M and that of NADPH was 3 mM, and the solution was incubated at 30 °C for 30 minutes. Then the solution was diluted with equal volume of
40 methanol, sufficiently stirred and centrifuged (15,000 x g, 5 min) to give the supernatant. The obtained supernatant was subjected to HPLC analysis according to the following conditions.

Microsomal fractions of yeasts expressing P450 2C19 showed good activity for the hydroxylation of the S-mephenytoin, while those of yeast cells expressing other human P450 molecular species showed no activity.

Conditions for HPLC

45

Column: COSMOSIL 5C18 (manufactured by Nakarai Tesq Co.)

Carrier: A:(Methanol)/(20 mM Potassium phosphate buffer, pH = 7.0) = 40/60

B:Methanol

50

Time (minute)	A/B
0-18	100/0
18-20	0/100
20-35	100/0

55

Temperature: room temperature

Detector: RI 14 C

Specimen amount: 100 μ l

Table 1. Results of the hydroxylation activity using human P450 molecular species

Human P450 molecular species											
Substrate	1A2	2C9	2E1	3A4	1A1	2A6	2B6	2C8	2C18	2C19	2D6
7-Ethoxycoumarin	++	+	++	+	+++	++	+++	+	+	+	++
Tolbutamide	-	+++	-	-	-	-	-	+	+	++	-
Testosterone	-	-	-	+++	+	-	-	+	-	-	-
Chlorzoxazone	++	+	+++	+	++	++	+	+	+	+	++
2-Aminoanthracene	+++	*	++	-	*	-	*	-	*	*	*
Acetanilide	+++	++	++	+	++	++	++	++	++	++	+++
Coumarin	-	-	-	-	-	+++	-	-	-	-	-
Debrisoquine	-	-	-	-	++	-	-	-	-	-	+++
S-Mephenytoin	-	-	-	-	-	-	-	-	-	+++	-

Hydroxylation activity is designated as follows: +, moderate activity; ++, good activity; +++, high activity; -, no activity; *, not examined.

55 Metabolism of chlorzoxazone using a mixture of microsomal fractions of transformed yeast cells

Microsomal fractions of yeast expressing cytochrome P450 prepared as above were mixed in the following molar ratios, and the hydroxylation activities of the mixed solutions were measured using

chlorzoxazone.

P450	System A	System B
3A4	35%	33%
2C9	25%	5.8%
2C8		5.8%
2C18		5.8%
2C19		5.8%
1A2	23%	19%
2E1	17%	15%
1A1		2.4%
2A6		3.0%
2B6		2.4%
2D6		2.4%

The substrate, [¹⁴C]chlorzoxazone and NADPH were added to the mixed yeast microsomal fractions, so that the concentrations of the compound and NADPH were 382 μM and 3 mM. The solutions were incubated at 37 °C for 30 min, and then 1 ml of dichloromethane was added thereto to stop the reaction. After stirring, dichloromethane layer was recovered by centrifugation (10,000 x g, 5 min). Then the solvent was evaporated by the stream of nitrogen gas. The obtained residue was dissolved in 54 μl of acetonitrile and 146 μl of water, the solution was subjected to HPLC analysis under the following conditions.

Conditions for HPLC

Column: COSMOSIL 5C18 (manufactured by Nakarai Tesq Co.)
Carrier: A(Acetonitrile/Water = 27/73)
B(Acetonitrile)

Time (minute)	A/B
0-15	100/0
15-17	0/100
17-25	100/0

Temperature: room temperature
Detector: RI ¹⁴C
Injection amount: 100 μl

The metabolites of chlorzoxazone observed by each of the mixed systems A and B were similar to those metabolites which Guengerich reported based on their experimental results by using human liver microsomal fractions (Guengerich, F.P., Chem. Toxicol., Vol.3, pp.566-573, 1990).

Furthermore, the metabolic turnover numbers were calculated for the human liver microsomal fraction (by Guengerich) and for the present yeast microsomal fractions.

The turnover numbers were calculated to be 1.8 and 1.6 in the mixed systems A and B, respectively. The turnover V for the human liver microsomal fraction was calculated using V_{max}, K_m and substrate concentration [S] described in the literature according to the following manner. The results are shown in Table 2. The values somewhat varied due to the difference of individuals, the lowest value being 1.0 and the highest value being 5.9. The values of V for the mixed system B and A fell within this range, both of which were the same level. It was confirmed that the four kinds of molecular species in system A can well reproduce the metabolic system in human liver in vitro.

A turnover V for human cytochrome P450 at an optional substrate concentration can be calculated by substituting V_{max} and K_m described in the literature and substrate concentration [S] of the present example into the Michaelis-Menten's equation:

$$V = (V_{\max} \cdot [S]) / (K_m + [S])$$

Table 2

Liver sample	Metabolic turnover V [product nmol/nmol P450/min]
#1001	5.9
KDL 14	2.2
KDL 21	1.7
KDL 23	3.0
KDL 27	5.0
H 10	1.1
H 11	1.0
H 12	4.2
H 13	3.3
H 14	2.1
H 15	4.3
H 16	4.0
H 17	3.6
H 18	3.4
Designations of the human liver sample were those used by Guengerich.	

Metabolism of debrisoquine using mixture of microsomal fractions of transformed yeast cells

Microsomal fractions of yeasts expressing human cytochrome P450 were mixed, and the hydroxylation activity of the mixed fraction was measured using debrisoquine. The mixing molar ratio of the human cytochrome P450 molecular species were as follows:

P450	Molar ratio
3A4	33%
2C9	5.8%
2C8	5.8%
2C18	5.8%
2C19	5.8%
1A2	19%
2E1	15%
1A1	2.4%
2B6	2.4%
2D6	2.4%

The substrate debrisoquine and NADPH were added to the mixed microsomal fraction solutions, so that the concentrations were 100 μ M for the NADPH and 6 mM for the compound. After the mixture was incubated at 37°C for 30 min, 50 μ l of 60% perchlorate was added to the solution to stop the reaction. The concentration of the perchlorate was finally 12.5% (v/v). After vigorous stirring, the mixture was centrifuged (15,000 \times g, 5 min) to recover the supernatant, which was subjected to HPLC analysis under the same conditions used for analyzing the metabolites of debrisoquine.

The metabolites well coincided with the metabolites which Kronbach reported based on the experiments to metabolize the debrisoquine using the human liver microsome (Methods in Enzymology, Vol.206, pp.509-517, 1991).

Metabolism of S-mephenytoin using mixture of microsomal fractions of transformed yeast cells

Microsomal fractions of yeasts expressing various human cytochrome P450 prepared were mixed, and the hydroxylation activity of the mixed fraction was measured for S-mephenytoin. The mixing ratio of the human cytochrome P450 molecular species was the same as that of the mixing system B as described above.

The substrate, [¹⁴C]S-mephenytoin and NADPH were added to the mixed microsomal fraction solutions, so that the concentrations were 28 μM for the NADPH and 6 mM for the compound. After the mixture was incubated at 37°C for 30 min, 250 μl of methanol was added to the solution to stop the reaction. After vigorous stirring, the mixture was centrifuged (15,000 x g, 5 min) to recover the supernatant, which was subjected to HPLC analysis under the same conditions used for the hydroxylation of S-mephenytoin using microsomal fraction. The metabolites obtained well coincided with the metabolites which Goldstein reported based on the experiments to metabolize the S-mephenytoin using the human liver microsome (Biochemistry, Vol.33, pp.1743-1752, 1994).

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT:
 (A) NAME: Sumitomo Chemical Company, Limited
 (B) STREET: 5-33, Kitahama 4-chome, Chuo-ku,
 (C) CITY: Osaka-shi, Osaka-fu
 (E) COUNTRY: Japan
 (F) POSTAL CODE (ZIP): none
- (ii) TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING HUMAN CYTOCHROME P450
- (iii) NUMBER OF SEQUENCES: 61
- (iv) COMPUTER READABLE FORM:
 (A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
- (vi) PRIOR APPLICATION DATA:
 (A) APPLICATION NUMBER: JP 201120/1993
 (B) FILING DATE: 20-JUL-1993
- (vi) PRIOR APPLICATION DATA:
 (A) APPLICATION NUMBER: JP 180246/1993
 (B) FILING DATE: 21-JUL-1993
- (vi) PRIOR APPLICATION DATA:
 (A) APPLICATION NUMBER: JP 208279/1993
 (B) FILING DATE: 30-JUL-1993

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1551 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1548
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG GCA TTG TCC CAG TCT GTT CCC TTC TCG GCC ACA GAG CTC CTC CTG
 Met Ala Leu Ser Gln Ser Val Pro Phe Ser Ala Thr Glu Leu Leu Leu
 1 5 10 15

48

	GCC TCT GCC ATC TTC TGC CTG GTA TTC TGG GTG CTC AAG GGT TTG AGG	95
	Ala Ser Ala Ile Phe Cys Leu Val Phe Trp Val Leu Lys Gly Leu Arg	
	20 25 30	
5	CCT CGG GTC CCC AAA GGC CTG AAA AGT CCA CCA GAG CCA TGG GGC TGG	144
	Pro Arg Val Pro Lys Gly Leu Lys Ser Pro Pro Glu Pro Trp Gly Trp	
	35 40 45	
	CCC TTG CTC GGG CAT GTG CTG ACC CTG GGG AAG AAC CCG CAC CTG GCA	192
10	Pro Leu Leu Gly His Val Leu Thr Leu Gly Lys Asn Pro His Leu Ala	
	50 55 60	
	CTG TCA AGG ATG AGC CAG CGC TAC GGG GAC GTC CTG CAG ATC CGC ATT	240
	Leu Ser Arg Met Ser Gln Arg Tyr Gly Asp Val Leu Gln Ile Arg Ile	
	65 70 75 80	
15	GGC TCC ACG CCC GTG CTG GTG CTG AGC CGC CTG GAC ACC ATC CGG CAG	288
	Gly Ser Thr Pro Val Leu Val Leu Ser Arg Leu Asp Thr Ile Arg Gln	
	85 90 95	
	GCC CTG GTG CGG CAG GGC GAC GAT TTC AAG GGC CGG CCT GAC CTC TAC	336
20	Ala Leu Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr	
	100 105 110	
	ACC TCC ACC CTC ATC ACT GAT GGC CAG AGC TTG ACC TTC AGC ACA GAC	384
	Thr Ser Thr Leu Ile Thr Asp Gly Gln Ser Leu Thr Phe Ser Thr Asp	
	115 120 125	
25	TCT GGA CCG GTG TGG GCT GCC CGC CGG CGC CTG GCC CAG AAT GCC CTC	432
	Ser Gly Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Ala Leu	
	130 135 140	
	AAC ACC TTC TCC ATC GCC TCT GAC CCA GCT TCC TCA TCC TCC TGC TAC	480
30	Asn Thr Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Ser Ser Cys Tyr	
	145 150 155 160	
	CTG GAG GAG CAT GTG AGC AAG GAG GCT AAG GCC CTG ATC AGC AGG TTG	528
	Leu Glu Glu His Val Ser Lys Glu Ala Lys Ala Leu Ile Ser Arg Leu	
	165 170 175	
35	CAG GAG CTG ATG GCA GGG CCT GGG CAC TTC GAC CCT TAC AAT CAG GTG	576
	Gln Glu Leu Met Ala Gly Pro Gly His Phe Asp Pro Tyr Asn Gln Val	
	180 185 190	
	GTG GTG TCA GTG GCC AAC GTC ATT GGT GCC ATG TGC TTC GGA CAG CAC	624
40	Val Val Ser Val Ala Asn Val Ile Gly Ala Met Cys Phe Gly Gln His	
	195 200 205	
	TTC CCT GAG AGT AGC GAT GAG ATG CTC AGC CTC GTG AAG AAC ACT CAT	672
	Phe Pro Glu Ser Ser Asp Glu Met Leu Ser Leu Val Lys Asn Thr His	
	210 215 220	
45	GAG TTC GTG GAG ACT GCC TCC TCC GGG AAC CCC CTG GAC TTC TTC CCC	720
	Glu Phe Val Glu Thr Ala Ser Ser Gly Asn Pro Leu Asp Phe Phe Pro	
	225 230 235 240	
	ATC CTT CGC TAC CTG CCT AAC CCT GCC CTG CAG AGG TTC AAG GCC TTC	768
50	Ile Leu Arg Tyr Leu Pro Asn Pro Ala Leu Gln Arg Phe Lys Ala Phe	
	245 250 255	

EP 0 644 267 A2

	AAC CAG AGG TTC CTG TGG TTC CTG CAG AAA ACA GTC CAG GAG CAC TAT	816
	Asn Gln Arg Phe Leu Trp Phe Leu Gln Lys Thr Val Gln Glu His Tyr	
	260 265 270	
5	CAG GAC TTT GAC AAG AAC AGT GTC CGG GAC ATC ACG GGT GCC CTG TTC	864
	Gln Asp Phe Asp Lys Asn Ser Val Arg Asp Ile Thr Gly Ala Leu Phe	
	275 280 285	
	AAG CAC AGC AAG AAG GGG CCT AGA GCC AGC GGC AAC CTC ATC CCA CAG	912
	Lys His Ser Lys Lys Gly Pro Arg Ala Ser Gly Asn Leu Ile Pro Gln	
	290 295 300	
10	GAG AAG ATT GTC AAC CTT GTC AAT GAC ATC TTT GGA GCA GGA TTT GAC	960
	Glu Lys Ile Val Asn Leu Val Asn Asp Ile Phe Gly Ala Gly Phe Asp	
	305 310 315 320	
15	ACA GTC ACC ACA GCC ATC TCC TGG AGC CTC ATG TAC CTT GTG ACC AAG	1008
	Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Thr Lys	
	325 330 335	
	CCT GAG ATA CAG AGG AAG ATC CAG AAG GAG CTG GAC ACT GTG ATT GGC	1056
	Pro Glu Ile Gln Arg Lys Ile Gln Lys Glu Leu Asp Thr Val Ile Gly	
	340 345 350	
20	AGG GAG CGG CGG CCC CGG CTC TCT GAC AGA CCC CAG CTG CCC TAC TTG	1104
	Arg Glu Arg Arg Pro Arg Leu Ser Asp Arg Pro Gln Leu Pro Tyr Leu	
	355 360 365	
25	GAG GCC TTC ATC CTG GAG ACC TTC CGA CAC TCC TCC TTC TTG CCC TTC	1152
	Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Leu Pro Phe	
	370 375 380	
	ACC ATC CCC CAC AGC ACA ACA AGG GAC ACA ACG CTG AAT GGC TTC TAC	1200
	Thr Ile Pro His Ser Thr Thr Arg Asp Thr Thr Leu Asn Gly Phe Tyr	
	385 390 395 400	
30	ATC CCC AAG AAA TGC TGT GTC TTC GTA AAC CAG TGG CAG GTC AAC CAT	1248
	Ile Pro Lys Lys Cys Cys Val Phe Val Asn Gln Trp Gln Val Asn His	
	405 410 415	
35	GAC CCA GAG CTG TGG GAG GAC CCC TCT GAG TTC CGG CCT GAG CGG TTC	1296
	Asp Pro Glu Leu Trp Glu Asp Pro Ser Glu Phe Arg Pro Glu Arg Phe	
	420 425 430	
	CTC ACC GCC GAT GGC ACT GCC ATT AAC AAG CCC TTG AGT GAG AAG ATG	1344
	Leu Thr Ala Asp Gly Thr Ala Ile Asn Lys Pro Leu Ser Glu Lys Met	
	435 440 445	
40	ATG CTG TTT GGC ATG GGT AAG CGC CGG TGT ATC GGG GAA GTC CTG GCC	1392
	Met Leu Phe Gly Met Gly Lys Arg Arg Cys Ile Gly Glu Val Leu Ala	
	450 455 460	
45	AAG TGG GAG ATC TTC CTC TTC CTG GCC ATC CTG CTA CAG CAA CTG GAG	1440
	Lys Trp Glu Ile Phe Leu Phe Leu Ala Ile Leu Leu Gln Gln Leu Glu	
	465 470 475 480	

50

55

TTC AGC GTG CCG CCG GGC GTG AAA GTC GAC CTG ACC CCC ATC TAC GGG 1488
 Phe Ser Val Pro Pro Gly Val Lys Val Asp Leu Thr Pro Ile Tyr Gly
 485 490 495
 5 CTG ACC ATG AAG CAC GCC CGC TGT GAA CAT GTC CAG GCG CGG CTG CGC 1536
 Leu Thr Met Lys His Ala Arg Cys Glu His Val Gln Ala Arg Leu Arg
 500 505 510
 TTC TCC ATC AAC TGA 1551
 Phe Ser Ile Asn
 515
 10

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 516 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 15

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Leu Ser Gln Ser Val Pro Phe Ser Ala Thr Glu Leu Leu Leu
 1 5 10 15
 Ala Ser Ala Ile Phe Cys Leu Val Phe Trp Val Leu Lys Gly Leu Arg
 20 25 30
 Pro Arg Val Pro Lys Gly Leu Lys Ser Pro Pro Glu Pro Trp Gly Trp
 35 40 45
 Pro Leu Leu Gly His Val Leu Thr Leu Gly Lys Asn Pro His Leu Ala
 50 55 60
 30 Leu Ser Arg Met Ser Gln Arg Tyr Gly Asp Val Leu Gln Ile Arg Ile
 65 70 75 80
 Gly Ser Thr Pro Val Leu Val Leu Ser Arg Leu Asp Thr Ile Arg Gln
 85 90 95
 35 Ala Leu Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr
 100 105 110
 Thr Ser Thr Leu Ile Thr Asp Gly Gln Ser Leu Thr Phe Ser Thr Asp
 115 120 125
 40 Ser Gly Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Ala Leu
 130 135 140
 Asn Thr Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Ser Cys Tyr
 145 150 155 160
 45 Leu Glu Glu His Val Ser Lys Glu Ala Lys Ala Leu Ile Ser Arg Leu
 165 170 175
 Gln Glu Leu Met Ala Gly Pro Gly His Phe Asp Pro Tyr Asn Gln Val
 180 185 190
 50

55

EP 0 644 267 A2

Val Val Ser Val Ala Asn Val Ile Gly Ala Met Cys Phe Gly Gln His
195 200 205

5 Phe Pro Glu Ser Ser Asp Glu Met Leu Ser Leu Val Lys Asn Thr His
210 215 220

Glu Phe Val Glu Thr Ala Ser Ser Gly Asn Pro Leu Asp Phe Phe Pro
225 230 235 240

10 Ile Leu Arg Tyr Leu Pro Asn Pro Ala Leu Gln Arg Phe Lys Ala Phe
245 250 255

Asn Gln Arg Phe Leu Trp Phe Leu Gln Lys Thr Val Gln Glu His Tyr
260 265 270

15 Gln Asp Phe Asp Lys Asn Ser Val Arg Asp Ile Thr Gly Ala Leu Phe
275 280 285

Lys His Ser Lys Lys Gly Pro Arg Ala Ser Gly Asn Leu Ile Pro Gln
290 295 300

20 Glu Lys Ile Val Asn Leu Val Asn Asp Ile Phe Gly Ala Gly Phe Asp
305 310 315 320

Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Thr Lys
325 330 335

25 Pro Glu Ile Gln Arg Lys Ile Gln Lys Glu Leu Asp Thr Val Ile Gly
340 345 350

Arg Glu Arg Arg Pro Arg Leu Ser Asp Arg Pro Gln Leu Pro Tyr Leu
355 360 365

30 Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Leu Pro Phe
370 375 380

Thr Ile Pro His Ser Thr Thr Arg Asp Thr Thr Leu Asn Gly Phe Tyr
385 390 395 400

Ile Pro Lys Lys Cys Cys Val Phe Val Asn Gln Trp Gln Val Asn His
405 410 415

40 Asp Pro Glu Leu Trp Glu Asp Pro Ser Glu Phe Arg Pro Glu Arg Phe
420 425 430

Leu Thr Ala Asp Gly Thr Ala Ile Asn Lys Pro Leu Ser Glu Lys Met
435 440 445

45 Met Leu Phe Gly Met Gly Lys Arg Arg Cys Ile Gly Glu Val Leu Ala
450 455 460

Lys Trp Glu Ile Phe Leu Phe Leu Ala Ile Leu Leu Gln Gln Leu Glu
465 470 475 480

50 Phe Ser Val Pro Pro Gly Val Lys Val Asp Leu Thr Pro Ile Tyr Gly
485 490 495

55

Leu Thr Met Lys His Ala Arg Cys Glu His Val Gln Ala Arg Leu Arg
 500 505 510

Phe Ser Ile Asn
 515

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1473 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1470

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

20	ATG GAT TCT ATT GTG TCC CTT GTG CTC TGT CTC TCA TGT TTG CTT CTC	48
	Met Asp Ser Ile Val Ser Leu Val Leu Cys Leu Ser Cys Leu Leu Leu	
	1 5 10 15	
	CTT TCA CTC TGG AGA CAG AGC TCT GGG AGA GGA AAA CTC CCT CCT GGC	96
	Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly	
	20 25 30	
25	CCC ACT CCT CTC CCA GTG ATT GGA AAT ATC CTA CAG ATA GGT ATT AAG	144
	Pro Thr Pro Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys	
	35 40 45	
30	GAC ATC AGC AAA TCC TTA ACC AAT CTC TCA AAG GTC TAT GGC CCT GTG	192
	Asp Ile Ser Lys Ser Leu Thr Asn Leu Ser Lys Val Tyr Gly Pro Val	
	50 55 60	
	TTC ACT CTG TAT TTT GGC CTG AAA CCC ATA GTG GTG CTG CAT GGA TAT	240
	Phe Thr Leu Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr	
	65 70 75 80	
35	GAA GCA GTG AAG GAA GCC CTG ATT GAT CTT GGA GAG GAG TTT TCT GGA	288
	Glu Ala Val Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly	
	85 90 95	
40	AGA GGC ATT TTC CCA CTG GCT GAA AGA GCT AAC AGA GGA TTT GGA ATT	336
	Arg Gly Ile Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile	
	100 105 110	
	GTT TTC AGC AAT GGA AAG AAA TGG AAG GAG ATC CGG CGT TTC TCC CTC	384
	Val Phe Ser Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu	
	115 120 125	
45	ATG ACG CTG CGG AAT TTT GGG ATG GGG AAG AGG AGC ATT GAG GAC CGT	432
	Met Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg	
	130 135 140	

50

55

EP 0 644 267 A2

	GTT CAA GAG GAA GCC CGC TGC CTT GTG GAG GAG TTG AGA AAA ACC AAG	480
	Val Gln Glu Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys	
	145 150 155 160	
5	GCC TCA CCC TGT GAT CCC ACT TTC ATC CTG GGC TGT GCT CCC TGC AAT	528
	Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn	
	165 170 175	
	GTG ATC TGC TCC ATT ATT TTC CAT AAA CGT TTT GAT TAT AAA GAT CAG	576
10	Val Ile Cys Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln	
	180 185 190	
	CAA TTT CTT AAC TTA ATG GAA AAG TTG AAT GAA AAC ATC AAG ATT TTG	624
	Gln Phe Leu Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Lys Ile Leu	
	195 200 205	
15	AGC AGC CCC TGG ATC CAG ATC TGC AAT AAT TTT TCT CCT ATC ATT GAT	672
	Ser Ser Pro Trp Ile Gln Ile Cys Asn Asn Phe Ser Pro Ile Ile Asp	
	210 215 220	
	TAC TTC CCG GGA ACT CAC AAC AAA TTA CTT AAA AAC GTT GCT TTT ATG	720
20	Tyr Phe Pro Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met	
	225 230 235 240	
	AAA AGT TAT ATT TTG GAA AAA GTA AAA GAA CAC CAA GAA TCA ATG GAC	768
	Lys Ser Tyr Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp	
	245 250 255	
25	ATG AAC AAC CCT CAG GAC TTT ATT GAT TGC TTC CTG ATG AAA ATG GAG	816
	Met Asn Asn Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu	
	260 265 270	
	AAG GAA AAG CAC AAC CAA CCA TCT GAA TTT ACT ATT GAA AGC TTG GAA	864
30	Lys Glu Lys His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu	
	275 280 285	
	AAC ACT GCA GTT GAC TTG TTT GGA GCT GGG ACA GAG ACG ACA AGC ACA	912
	Asn Thr Ala Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr	
	290 295 300	
35	ACC CTG AGA TAT GCT CTC CTT CTC CTG CTG AAG CAC CCA GAG GTC ACA	960
	Thr Leu Arg Tyr Ala Leu Leu Leu Leu Lys His Pro Glu Val Thr	
	305 310 315 320	
	GCT AAA GTC CAG GAA GAG ATT GAA CGT GTG ATT GGC AGA AAC CGG AGC	1008
40	Ala Lys Val Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser	
	325 330 335	
	CCC TGC ATG CAA GAC AGG AGC CAC ATG CCC TAC ACA GAT GCT GTG GTG	1056
	Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val	
	340 345 350	
45	CAC GAG GTC CAG AGA TAC ATT GAC CTT CTC CCC ACC AGC CTG CCC CAT	1104
	His Glu Val Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His	
	355 360 365	
	GCA GTG ACC TGT GAC ATT AAA TTC AGA AAC TAT CTC ATT CCC AAG GGC	1152
50	Ala Val Thr Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly	
	370 375 380	

ACA ACC ATA TTA ATT TCC CTG ACT TCT GTG CTA CAT GAC AAC AAA GAA 1200
 Thr Thr Ile Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu
 385 390 395 400
 5 TTT CCC AAC CCA GAG ATG TTT GAC CCT CAT CAC TTT CTG GAT GAA GGT 1248
 Phe Pro Asn Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly
 405 410 415
 GGC AAT TTT AAG AAA AGT AAA TAC TTC ATG CCT TTC TCA GCA GGA AAA 1296
 Gly Asn Phe Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys
 10 420 425 430
 CGG ATT TGT GTG GGA GAA GCC CTG GCC GGC ATG GAG CTG TTT TTA TTC 1344
 Arg Ile Cys Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe
 435 440 445
 15 CTG ACC TCC ATT TTA CAG AAC TTT AAC CTG AAA TCT CTG GTT GAC CCA 1392
 Leu Thr Ser Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro
 450 455 460
 AAG AAC CTT GAG ACC ACT CCA GTT GTC AAT GGA TTT GCC TCT GTG CCG 1440
 Lys Asn Leu Asp Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro
 20 465 470 475 480
 CCC TTC TAC CAG CTG TGC TTC ATT CCT GTC TGA 1473
 Pro Phe Tyr Gln Leu Cys Phe Ile Pro Val
 485 490

25 (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 490 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

35 Met Asp Ser Ile Val Ser Leu Val Leu Cys Leu Ser Cys Leu Leu Leu
 1 5 10 15
 Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly
 20 25 30
 40 Pro Thr Pro Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Gly Ile Lys
 35 40 45
 Asp Ile Ser Lys Ser Leu Thr Asn Leu Ser Lys Val Tyr Gly Pro Val
 50 55 60
 45 Phe Thr Leu Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr
 65 70 75 80
 Glu Ala Val Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly
 85 90 95

50

55

Arg Gly Ile Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile
 100 105 110
 5 Val Phe Ser Asn Gly Lys Lys Trp Lys Glu Ile Arg Arg Phe Ser Leu
 115 120 125
 Met Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg
 130 135 140
 10 Val Gln Glu Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys
 145 150 155 160
 Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
 165 170 175
 15 Val Ile Cys Ser Ile Ile Phe His Lys Arg Phe Asp Tyr Lys Asp Gln
 180 185 190
 Gln Phe Leu Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Lys Ile Leu
 195 200 205
 20 Ser Ser Pro Trp Ile Gln Ile Cys Asn Asn Phe Ser Pro Ile Ile Asp
 210 215 220
 Tyr Phe Pro Gly Thr His Asn Lys Leu Leu Lys Asn Val Ala Phe Met
 225 230 235 240
 25 Lys Ser Tyr Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp
 245 250 255
 Met Asn Asn Pro Gln Asp Phe Ile Asp Cys Phe Leu Met Lys Met Glu
 260 265 270
 30 Lys Glu Lys His Asn Gln Pro Ser Glu Phe Thr Ile Glu Ser Leu Glu
 275 280 285
 Asn Thr Ala Val Asp Leu Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr
 290 295 300
 35 Thr Leu Arg Tyr Ala Leu Leu Leu Leu Lys His Pro Glu Val Thr
 305 310 315 320
 Ala Lys Val Gln Glu Glu Ile Glu Arg Val Ile Gly Arg Asn Arg Ser
 325 330 335
 40 Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val
 340 345 350
 His Glu Val Gln Arg Tyr Ile Asp Leu Leu Pro Thr Ser Leu Pro His
 355 360 365
 45 Ala Val Thr Cys Asp Ile Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly
 370 375 380
 Thr Thr Ile Leu Ile Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu
 385 390 395 400
 50 Phe Pro Asn Pro Glu Met Phe Asp Pro His His Phe Leu Asp Glu Gly
 405 410 415

Gly Asn Phe Lys Lys Ser Lys Tyr Phe Met Pro Phe Ser Ala Gly Lys
 420 425 430
 Arg Ile Cys Val Gly Glu Ala Leu Ala Gly Met Glu Leu Phe Leu Phe
 5 435 440 445
 Leu Thr Ser Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Val Asp Pro
 450 455 460
 Lys Asn Leu Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro
 10 465 470 475 480
 Pro Phe Tyr Gln Leu Cys Phe Ile Pro Val
 485 490

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1482 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1479

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATG TCT GCC CTC GGA GTC ACC GTG GCC CTG CTG GTG TGG GCG GCC TTC 48
 Met Ser Ala Leu Gly Val Thr Val Ala Leu Val Trp Ala Ala Phe
 1 5 10 15
 CTC CTG CTG GTG TCC ATG TGG AGG CAG GTG CAC AGC AGC TGG AAT CTG 96
 Leu Leu Leu Val Ser Met Trp Arg Gln Val His Ser Ser Trp Asn Leu
 20 25 30
 CCC CCA GGC CCT TTC CCG CTT CCC ATC ATC GGG AAC CTC TTC CAG TTG 144
 Pro Pro Gly Pro Phe Pro Leu Pro Ile Ile Gly Asn Leu Phe Gln Leu
 35 40 45
 GAA TTG AAG AAT ATT CCC AAG TCC TTC ACC CGG TTG GCC CAG CGC TTC 192
 Glu Leu Lys Asn Ile Pro Lys Ser Phe Thr Arg Leu Ala Gln Arg Phe
 50 55 60
 GGG CCG GTG TTC ACG CTG TAC GTG GGC TCG CAG CGC ATG GTG GTG ATG 240
 Gly Pro Val Phe Thr Leu Tyr Val Gly Ser Gln Arg Met Val Val Met
 65 70 75 80
 CAC GGC TAC AAG GCG GTG AAG GAA GCG CTG GAC TAC AAG GAC GAG 288
 His Gly Tyr Lys Ala Val Lys Glu Ala Leu Asp Tyr Lys Asp Glu
 85 90 95
 TTC TCG GGC AGA GGC GAC CTC CCC GCG TTC CAT GCG CAC AGG GAC AGG 336
 Phe Ser Gly Arg Gly Asp Leu Pro Ala Phe His Ala His Arg Asp Arg
 100 105 110

EP 0 644 267 A2

	GGA ATC ATT TTT AAT AAT GGA CCT ACC TGG AAG GAC ATC CGG CGG TTT	384
	Gly Ile Ile Phe Asn Asn Gly Pro Thr Trp Lys Asp Ile Arg Arg Phe	
	115 120 125	
5	TCC CTG ACC ACC CTC CGG AAC TAT GGG ATG GGG AAA CAG GGC AAT GAG	432
	Ser Leu Thr Thr Leu Arg Asn Tyr Gly Met Gly Lys Gln Gly Asn Glu	
	130 135 140	
10	AGC CGG ATC CAG AGG GAG GCC CAC TTC CTG CTG GAA GCA CTC AGG AAG	480
	Ser Arg Ile Gln Arg Glu Ala His Phe Leu Leu Glu Ala Leu Arg Lys	
	145 150 155 160	
	ACC CAA GGC CAG CCT TTC GAC CCC ACC TTC CTC ATC GGG TGC GCG CCC	528
	Thr Gln Gly Gln Pro Phe Asp Pro Thr Phe Leu Ile Gly Cys Ala Pro	
	165 170 175	
15	TGC AAC GTC ATA GCC GAC ATC CTC TTC CGC AAG CAT TTT GAC TAC AAT	576
	Cys Asn Val Ile Ala Asp Ile Leu Phe Arg Lys His Phe Asp Tyr Asn	
	180 185 190	
20	GAT GAG AAG TTT CTA AGG CTG ATG TAT TTG TTT AAT GAG AAC TTC CAC	624
	Asp Glu Lys Phe Leu Arg Leu Met Tyr Leu Phe Asn Glu Asn Phe His	
	195 200 205	
	CTA CTC AGC ACT CCC TGG CTC CAG CTT TAC AAT AAT TTT CCC AGC TTT	672
	Leu Leu Ser Thr Pro Trp Leu Gln Leu Tyr Asn Asn Phe Pro Ser Phe	
	210 215 220	
25	CTA CAC TAC TTG CCT GGA AGC CAC AGA AAA GTC ATA AAA AAT GTG GCT	720
	Leu His Tyr Leu Pro Gly Ser His Arg Lys Val Ile Lys Asn Val Ala	
	225 230 235 240	
30	GAA GTA AAA GAG TAT GTG TCT GAA AGG GTG AAG GAG CAC CAT CAA TCT	768
	Glu Val Lys Glu Tyr Val Ser Glu Arg Val Lys Glu His His Gln Ser	
	245 250 255	
	CTG GAC CCC AAC TGT CCC CGG GAC CTC ACC GAC TGC CTG CTC GTG GAA	816
	Leu Asp Pro Asn Cys Pro Arg Asp Leu Thr Asp Cys Leu Leu Val Glu	
	260 265 270	
35	ATG GAG AAG GAA AAG CAC AGT GCA GAG CGC TTG TAC ACA ATG GAC GGT	864
	Met Glu Lys Glu Lys His Ser Ala Glu Arg Leu Tyr Thr Met Asp Gly	
	275 280 285	
40	ATC ACC GTG ACT GTG GCC GAC CTG TTC TTT GCG GGG ACA GAG ACC ACC	912
	Ile Thr Val Thr Val Ala Asp Leu Phe Phe Ala Gly Thr Glu Thr Thr	
	290 295 300	
	AGC ACA ACT CTG AGA TAT GGG CTC CTG ATT CTC ATG AAA TAC CCT GAG	960
	Ser Thr Thr Leu Arg Tyr Gly Leu Leu Ile Leu Met Lys Tyr Pro Glu	
	305 310 315 320	
45	ATC GAA GAG AAG CTC CAT GAA GAA ATT GAC AGG GTG ATT GGG CCA AGC	1008
	Ile Glu Glu Lys Leu His Glu Glu Ile Asp Arg Val Ile Gly Pro Ser	
	325 330 335	

50

55

	CGA ATC CCT GCC ATC AAG GAT AGG CAA GAG ATG CCC TAC ATG GAT GCT	1056
	Arg Ile Pro Ala Ile Lys Asp Arg Gln Glu Met Pro Tyr Met Asp Ala	
	340 345 350	
5	GTG GTG CAT GAG ATT CAG CGG TTC ATC ACC CTC GTG CCC TCC AAC CTG	1104
	Val Val His Glu Ile Gln Arg Phe Ile Thr Leu Val Pro Ser Asn Leu	
	355 360 365	
	CCC CAT GAA GCA ACC CGA GAC ACC ATT TTC AGA GGA TAC CTC ATC CCC	1152
10	Pro His Glu Ala Thr Arg Asp Thr Ile Phe Arg Gly Tyr Leu Ile Pro	
	370 375 380	
	AAG GGC ACA GTC GTA GTG CCA ACT CTG GAC TCT GTT TTG TAT GAC AAC	1200
	Lys Gly Thr Val Val Pro Thr Leu Asp Ser Val Leu Tyr Asp Asn	
	385 390 395 400	
15	CAA GAA TTT CCT GAT CCA GAA AAG TTT AAG CCA GAA CAC TTC CTG AAT	1248
	Gln Glu Phe Pro Asp Pro Glu Lys Phe Lys Pro Glu His Phe Leu Asn	
	405 410 415	
	GAA AAT GGA AAG TTC AAG TAC AGT GAC TAT TTC AAG CCA TTT TCC ACA	1296
20	Glu Asn Gly Lys Phe Lys Tyr Ser Asp Tyr Phe Lys Pro Phe Ser Thr	
	420 425 430	
	GGA AAA CGA GTG TGT GCT GGA GAA GGC CTG GCT CGC ATG GAG TTG TTT	1344
	Gly Lys Arg Val Cys Ala Gly Glu Gly Leu Ala Arg Met Glu Leu Phe	
	435 440 445	
25	CTT TTG TTG TGT GCC ATT TTG CAG CAT TTT AAT TTG AAG CCT CTC GTT	1392
	Leu Leu Leu Cys Ala Ile Leu Gln His Phe Asn Leu Lys Pro Leu Val	
	450 455 460	
	GAC CCA AAG GAT ATC GAC CTC AGC CCT ATA CAT ATT GGG TTT GGC TGT	1440
30	Asp Pro Lys Asp Ile Asp Leu Ser Pro Ile His Ile Gly Phe Gly Cys	
	465 470 475 480	
	ATC CCA CCA CGT TAC AAA CTC TGT GTC ATT CCC CGC TCA TGA	1482
	Ile Pro Pro Arg Tyr Lys Leu Cys Val Ile Pro Arg Ser	
	485 490	

35 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 493 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

45	Met Ser Ala Leu Gly Val Thr Val Ala Leu Leu Val Trp Ala Ala Phe
	1 5 10 15
	Leu Leu Leu Val Ser Met Trp Arg Gln Val His Ser Ser Trp Asn Leu
	20 25 30

50

55

Pro Pro Gly Pro Phe Pro Leu Pro Ile Ile Gly Asn Leu Phe Gln Leu
 35 40 45
 Glu Leu Lys Asn Ile Pro Lys Ser Phe Thr Arg Leu Ala Gln Arg Phe
 5 50 55 60
 Gly Pro Val Phe Thr Leu Tyr Val Gly Ser Gln Arg Met Val Val Met
 65 70 75 80
 His Gly Tyr Lys Ala Val Lys Glu Ala Leu Leu Asp Tyr Lys Asp Glu
 10 85 90 95
 Phe Ser Gly Arg Gly Asp Leu Pro Ala Phe His Ala His Arg Asp Arg
 100 105 110
 Gly Ile Ile Phe Asn Asn Gly Pro Thr Trp Lys Asp Ile Arg Arg Phe
 115 115 120 125
 Ser Leu Thr Thr Leu Arg Asn Tyr Gly Met Gly Lys Gln Gly Asn Glu
 130 135 140
 Ser Arg Ile Gln Arg Glu Ala His Phe Leu Leu Glu Ala Leu Arg Lys
 145 150 155 160
 Thr Gln Gly Gln Pro Phe Asp Pro Thr Phe Leu Ile Gly Cys Ala Pro
 165 170 175
 Cys Asn Val Ile Ala Asp Ile Leu Phe Arg Lys His Phe Asp Tyr Asn
 180 185 190
 Asp Glu Lys Phe Leu Arg Leu Met Tyr Leu Phe Asn Glu Asn Phe His
 195 200 205
 Leu Leu Ser Thr Pro Trp Leu Gln Leu Tyr Asn Asn Phe Pro Ser Phe
 210 215 220
 Leu His Tyr Leu Pro Gly Ser His Arg Lys Val Ile Lys Asn Val Ala
 225 230 235 240
 Glu Val Lys Glu Tyr Val Ser Glu Arg Val Lys Glu His His Gln Ser
 245 250 255
 Leu Asp Pro Asn Cys Pro Arg Asp Leu Thr Asp Cys Leu Leu Val Glu
 260 265 270
 Met Glu Lys Glu Lys His Ser Ala Glu Arg Leu Tyr Thr Met Asp Gly
 275 280 285
 Ile Thr Val Thr Val Ala Asp Leu Phe Phe Ala Gly Thr Glu Thr Thr
 290 295 300
 Ser Thr Thr Leu Arg Tyr Gly Leu Leu Ile Leu Met Lys Tyr Pro Glu
 305 310 315 320
 Ile Glu Glu Lys Leu His Glu Glu Ile Asp Arg Val Ile Gly Pro Ser
 325 330 335
 Arg Ile Pro Ala Ile Lys Asp Arg Gln Glu Met Pro Tyr Met Asp Ala
 340 345 350

Val Val His Glu Ile Gln Arg Phe Ile Thr Leu Val Pro Ser Asn Leu
355 360 365

5 Pro His Glu Ala Thr Arg Asp Thr Ile Phe Arg Gly Tyr Leu Ile Pro
370 375 380

Lys Gly Thr Val Val Val Pro Thr Leu Asp Ser Val Leu Tyr Asp Asn
385 390 395 400

10 Gln Glu Phe Pro Asp Pro Glu Lys Phe Lys Pro Glu His Phe Leu Asn
405 410 415

Glu Asn Gly Lys Phe Lys Tyr Ser Asp Tyr Phe Lys Pro Phe Ser Thr
420 425 430

15 Gly Lys Arg Val Cys Ala Gly Glu Gly Leu Ala Arg Met Glu Leu Phe
435 440 445

Leu Leu Leu Cys Ala Ile Leu Gln His Phe Asn Leu Lys Pro Leu Val
450 455 460

20 Asp Pro Lys Asp Ile Asp Leu Ser Pro Ile His Ile Gly Phe Gly Cys
465 470 475 480

Ile Pro Pro Arg Tyr Lys Leu Cys Val Ile Pro Arg Ser
485 490

(2) INFORMATION FOR SEQ ID NO: 7:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1512 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30 (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1509

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATG GCT CTC ATC CCA GAC TTG GCC ATG GAA ACC TGG CTT CTC CTG GCT	48
Met Ala Leu Ile Pro Asp Leu Ala Met Glu Thr Trp Leu Leu Leu Ala	
1 5 10 15	
40 GTC AGC CTG GTG CTC CTC TAT CTA TAT GGA ACC CAT TCA CAT GGA CTT	96
Val Ser Leu Val Leu Leu Tyr Leu Tyr Gly Thr His Ser His Gly Leu	
20 25 30	
144 TTT AAG AAG CTT GGA ATT CCA GGG CCC ACA CCT CTG CCT TTT TTG GGA	
Phe Lys Lys Leu Gly Ile Pro Gly Pro Thr Pro Leu Pro Phe Leu Gly	
35 40 45	
192 AAT ATT TTG TCC TAC CAT AAG GGC TTT TGT ATG TTT GAC ATG GAA TGT	
Asn Ile Leu Ser Tyr His Lys Gly Phe Cys Met Phe Asp Met Glu Cys	
50 55 60	

50

55

	CAT	AAA	AAG	TAT	GGA	AAA	GTG	TGG	GGC	TTT	TAT	GAT	GGT	CAA	CAG	CCT	240
	His	Lys	Lys	Tyr	Gly	Lys	Val	Trp	Gly	Phe	Tyr	Asp	Gly	Gln	Gln	Pro	
	65					70					75					80	
5	GTG	CTG	GCT	ATC	ACA	GAT	CCT	GAC	ATG	ATC	AAA	ACA	GTG	CTA	GTG	AAA	288
	Val	Leu	Ala	Ile	Thr	Asp	Pro	Asp	Met	Ile	Lys	Thr	Val	Leu	Val	Lys	
					85					90					95		
	GAA	TGT	TAT	TCT	GTC	TTC	ACA	AAC	CGG	AGG	CCT	TTT	GGT	CCA	GTG	GGA	336
10	Glu	Cys	Tyr		Val	Phe	Thr	Asn	Arg	Arg	Pro	Phe	Gly	Pro	Val	Gly	
				100					105					110			
	TTT	ATG	AAA	AGT	GCC	ATC	TCT	ATA	GCT	GAG	GAT	GAA	GAA	TGG	AAG	AGA	384
	Phe	Met	Lys	Ser	Ala	Ile	Ser	Ile	Ala	Glu	Asp	Glu	Glu	Trp	Lys	Arg	
			115					120					125				
15	TTA	CGA	TCA	TTG	CTG	TCT	CCA	ACC	TTC	ACC	AGT	GGA	AAA	CTC	AAG	GAG	432
	Leu	Arg	Ser	Leu	Leu	Ser		Thr	Phe	Thr	Ser	Gly	Lys	Leu	Lys	Glu	
		130					135					140					
	ATG	GTC	CCT	ATC	ATT	GCC	CAG	TAT	GGA	GAT	GTG	TTG	GTG	AGA	AAT	CTG	480
20	Met	Val	Pro	Ile	Ile	Ala	Gln	Tyr	Gly	Asp	Val	Leu	Val	Arg	Asn	Leu	
	145					150				155						160	
	AGG	CGG	GAA	GCA	GAG	ACA	GGC	AAG	CCT	GTC	ACC	TTG	AAA	GAC	GTC	TTT	528
	Arg	Arg	Glu	Ala	Glu	Thr	Gly	Lys	Pro	Val	Thr	Leu	Lys	Asp	Val	Phe	
					165					170					175		
25	GGG	GCC	TAC	AGC	ATG	GAT	GTG	ATC	ACT	AGC	ACA	TCA	TTT	GGA	GTG	AAC	576
	Gly	Ala	Tyr	Ser	Met	Asp	Val	Ile	Thr	Ser	Thr	Ser	Phe	Gly	Val	Asn	
				180					185					190			
	ATC	GAC	TCT	CTC	AAC	AAT	CCA	CAA	GAC	CCC	TTT	GTG	GAA	AAC	ACC	AAG	624
30	Ile	Asp	Ser	Leu	Asn	Asn	Pro	Gln	Asp	Pro	Phe	Val	Glu	Asn	Thr	Lys	
			195				200						205				
	AAG	CTT	TTA	AGA	TTT	GAT	TTT	TTG	GAT	CCA	TTC	TTT	CTC	TCA	ATA	ACA	672
	Lys	Leu	Leu	Arg	Phe	Asp	Phe	Leu	Asp	Pro	Phe	Phe	Leu	Ser	Ile	Thr	
		210				215						220					
35	GTC	TTT	CCA	TTC	CTC	ATC	CCA	ATT	CTT	GAA	GTA	TTA	AAT	ATC	TGT	GTG	720
	Val	Phe	Pro	Phe	Leu	Ile	Pro	Ile	Leu	Glu	Val	Leu	Asn	Ile	Cys	Val	
	225					230				235						240	
	TTT	CCA	AGA	GAA	GTT	ACA	AAT	TTT	TTA	AGA	AAA	TCT	GTA	AAA	AGG	ATG	768
40	Phe	Pro	Arg	Glu	Val	Thr	Asn	Phe	Leu	Arg	Lys	Ser	Val	Lys	Arg	Met	
					245					250					255		
	AAA	GAA	AGT	CGC	CTC	GAA	GAT	ACA	CAA	AAG	CAC	CGA	GTG	GAT	TTC	CTT	816
	Lys	Glu	Ser	Arg	Leu	Glu	Asp	Thr	Gln	Lys	His	Arg	Val	Asp	Phe	Leu	
				260					265					270			
45	CAG	CTG	ATG	ATT	GAC	TCT	CAG	AAT	TCA	AAA	GAA	ACT	GAG	TCC	CAC	AAA	864
	Gln	Leu	Met	Ile	Asp	Ser	Gln	Asn	Ser	Lys	Glu	Thr	Glu	Ser	His	Lys	
			275					280					285				

50

55

EP 0 644 267 A2

	GCT CTG TCC GAT CTG GAG CTC GTG GCC CAA TCA ATT ATC TTT ATT TTT	912
	Ala Leu Ser Asp Leu Glu Leu Val Ala Gln Ser Ile Ile Phe Ile Phe	
	290 295 300	
5	GCT GGC TAT GAA ACC ACG AGC AGT GTT CTC TCC TTC ATT ATG TAT GAA	960
	Ala Gly Tyr Glu Thr Thr Ser Ser Val Leu Ser Phe Ile Met Tyr Glu	
	305 310 315 320	
	CTG GCC ACT CAC CCT GAT GTC CAG CAG AAA CTG CAG GAG GAA ATT GAT	1008
	Leu Ala Thr His Pro Asp Val Gln Gln Lys Leu Gln Glu Glu Ile Asp	
	325 330 335	
10	GCA GTT TTA CCC AAT AAG GCA CCA CCC ACC TAT GAT ACT GTG CTA CAG	1056
	Ala Val Leu Pro Asn Lys Ala Pro Pro Thr Tyr Asp Thr Val Leu Gln	
	340 345 350	
	ATG GAG TAT CTT GAC ATG GTG GTG AAT GAA ACG CTC AGA TTA TTC CCA	1104
15	Met Glu Tyr Leu Asp Met Val Val Asn Glu Thr Leu Arg Leu Phe Pro	
	355 360 365	
	ATT GCT ATG AGA CTT GAG AGG GTC TGC AAA AAA GAT GTT GAG ATC AAT	1152
	Ile Ala Met Arg Leu Glu Arg Val Cys Lys Lys Asp Val Glu Ile Asn	
	370 375 380	
20	GGG ATG TTC ATT CCC AAA GGG TGG GTG GTG ATG ATT CCA AGC TAT GCT	1200
	Gly Met Phe Ile Pro Lys Gly Trp Val Val Met Ile Pro Ser Tyr Ala	
	385 390 395 400	
	CTT CAC CGT GAC CCA AAG TAC TGG ACA GAG CCT GAG AAG TTC CTC CCT	1248
25	Leu His Arg Asp Pro Lys Tyr Trp Thr Glu Pro Glu Lys Phe Leu Pro	
	405 410 415	
	GAA AGA TTC AGC AAG AAG AAC AAG GAC AAC ATA GAT CCT TAC ATA TAC	1296
	Glu Arg Phe Ser Lys Lys Asn Lys Asp Asn Ile Asp Pro Tyr Ile Tyr	
	420 425 430	
30	ACA CCC TTT GGA AGT GGA CCC AGA AAC TGC ATT GGC ATG AGG TTT GCT	1344
	Thr Pro Phe Gly Ser Gly Pro Arg Asn Cys Ile Gly Met Arg Phe Ala	
	435 440 445	
	CTC ATG AAC ATG AAA CTT GCT CTA ATC AGA GTC CTT CAG AAC TTC TCC	1392
35	Leu Met Asn Met Lys Leu Ala Leu Ile Arg Val Leu Gln Asn Phe Ser	
	450 455 460	
	TTC AAA CCT TGT AAA GAA ACA CAG ATC CCC CTG AAA TTA AGC TTA GGA	1440
	Phe Lys Pro Cys Lys Glu Thr Gln Ile Pro Leu Lys Leu Ser Leu Gly	
	465 470 475 480	
40	GGA CTT CTT CAA CCA GAA AAA CCC GTT GTT CTA AAG GTT GAG TCA AGG	1488
	Gly Leu Leu Gln Pro Glu Lys Pro Val Val Leu Lys Val Glu Ser Arg	
	485 490 495	
	GAT GGC ACC GTA AGT GGA GCC TGA	1512
45	Asp Gly Thr Val Ser Gly Ala	
	500	

50

55

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 503 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

5 Met Ala Leu Ile Pro Asp Leu Ala Met Glu Thr Trp Leu Leu Leu Ala
 1 5 10 15
 Val Ser Leu Val Leu Leu Tyr Leu Tyr Gly Thr His Ser His Gly Leu
 20 25 30
 Phe Lys Lys Leu Gly Ile Pro Gly Pro Thr Pro Leu Pro Phe Leu Gly
 35 40 45
 Asn Ile Leu Ser Tyr His Lys Gly Phe Cys Met Phe Asp Met Glu Cys
 50 55 60
 20 His Lys Lys Tyr Gly Lys Val Trp Gly Phe Tyr Asp Gly Gln Gln Pro
 65 70 75 80
 Val Leu Ala Ile Thr Asp Pro Asp Met Ile Lys Thr Val Leu Val Lys
 85 90 95
 Glu Cys Tyr Ser Val Phe Thr Asn Arg Arg Pro Phe Gly Pro Val Gly
 100 105 110
 30 Phe Met Lys Ser Ala Ile Ser Ile Ala Glu Asp Glu Glu Trp Lys Arg
 115 120 125
 Leu Arg Ser Leu Leu Ser Pro Thr Phe Thr Ser Gly Lys Leu Lys Glu
 130 135 140
 35 Met Val Pro Ile Ile Ala Gln Tyr Gly Asp Val Leu Val Arg Asn Leu
 145 150 155 160
 Arg Arg Glu Ala Glu Thr Gly Lys Pro Val Thr Leu Lys Asp Val Phe
 165 170 175
 40 Gly Ala Tyr Ser Met Asp Val Ile Thr Ser Thr Ser Phe Gly Val Asn
 180 185 190
 Ile Asp Ser Leu Asn Asn Pro Gln Asp Pro Phe Val Glu Asn Thr Lys
 195 200 205
 45 Lys Leu Leu Arg Phe Asp Phe Leu Asp Pro Phe Phe Leu Ser Ile Thr
 210 215 220
 Val Phe Pro Phe Leu Ile Pro Ile Leu Glu Val Leu Asn Ile Cys Val
 225 230 235 240
 50 Phe Pro Arg Glu Val Thr Asn Phe Leu Arg Lys Ser Val Lys Arg Met
 245 250 255

55

Lys Glu Ser Arg Leu Glu Asp Thr Gln Lys His Arg Val Asp Phe Leu
 260 265 270
 5 Gln Leu Met Ile Asp Ser Gln Asn Ser Lys Glu Thr Glu Ser His Lys
 275 280 285
 Ala Leu Ser Asp Leu Glu Leu Val Ala Gln Ser Ile Ile Phe Ile Phe
 290 295 300
 10 Ala Gly Tyr Glu Thr Thr Ser Ser Val Leu Ser Phe Ile Met Tyr Glu
 305 310 315 320
 Leu Ala Thr His Pro Asp Val Gln Gln Lys Leu Gln Glu Glu Ile Asp
 325 330 335
 15 Ala Val Leu Pro Asn Lys Ala Pro Pro Thr Tyr Asp Thr Val Leu Gln
 340 345 350
 Met Glu Tyr Leu Asp Met Val Val Asn Glu Thr Leu Arg Leu Phe Pro
 355 360 365
 20 Ile Ala Met Arg Leu Glu Arg Val Cys Lys Lys Asp Val Glu Ile Asn
 370 375 380
 Gly Met Phe Ile Pro Lys Gly Trp Val Val Met Ile Pro Ser Tyr Ala
 385 390 395 400
 25 Leu His Arg Asp Pro Lys Tyr Trp Thr Glu Pro Glu Lys Phe Leu Pro
 405 410 415
 Glu Arg Phe Ser Lys Lys Asn Lys Asp Asn Ile Asp Pro Tyr Ile Tyr
 420 425 430
 Thr Pro Phe Gly Ser Gly Pro Arg Asn Cys Ile Gly Met Arg Phe Ala
 435 440 445
 35 Leu Met Asn Met Lys Leu Ala Leu Ile Arg Val Leu Gln Asn Phe Ser
 450 455 460
 Phe Lys Pro Cys Lys Glu Thr Gln Ile Pro Leu Lys Leu Ser Leu Gly
 465 470 475 480
 40 Gly Leu Leu Gln Pro Glu Lys Pro Val Val Leu Lys Val Glu Ser Arg
 485 490 495
 Asp Gly Thr Val Ser Gly Ala
 500

45

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50

55

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1536

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

1	ATG CTT TTC CCA ATC TCC ATG TCG GCC ACG GAG TTT CTT CTG GCC TCT	48
	Met Leu Phe Pro Ile Ser Met Ser Ala Thr Glu Phe Leu Leu Ala Ser	
10	1 5 10 15	
20	GTC ATC TTC TGT CTG GTA TTC TGG GTA ATC AGG GCC TCA AGA CCT CAG	96
	Val Ile Phe Cys Leu Val Phe Trp Val Ile Arg Ala Ser Arg Pro Gln	
	20 25 30	
15	GTC CCC AAA GGC CTG AAG AAT CCA CCA GGG CCA TGG GGC TGG CCT CTG	144
	Val Pro Lys Gly Leu Lys Asn Pro Pro Gly Pro Trp Gly Trp Pro Leu	
	35 40 45	
20	ATT GGG CAC ATG CTG ACC CTG GGA AAG AAC CCG CAC CTG GCA CTG TCA	192
	Ile Gly His Met Leu Thr Leu Gly Lys Asn Pro His Leu Ala Leu Ser	
	50 55 60	
25	AGG ATG AGC CAG CAG TAT GGG GAC GTG CTG CAG ATC CGA ATT GGC TTC	240
	Arg Met Ser Gln Gln Tyr Gly Asp Val Leu Gln Ile Arg Ile Gly Ser	
	65 70 75 80	
30	ACA CCC GTG GTG GTG CTG AGC GGC CTG GAC ACC ATC CGG CAG GCC CTG	288
	Thr Pro Val Val Val Leu Ser Gly Leu Asp Thr Ile Arg Gln Ala Leu	
	85 90 95	
35	GTG CGG CAG GGC GAT GAT TTC AAG GGC CGG CCC GAC CTC TAC ACC TTC	336
	Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr Thr Phe	
	100 105 110	
40	ACC CTC ATC AGT AAT GGT CAG AGC ATG TCC TTC AGC CCA GAC TCT GGA	384
	Thr Leu Ile Ser Asn Gly Gln Ser Met Ser Phe Ser Pro Asp Ser Gly	
	115 120 125	
45	CCA GTG TGG GCT GCC CGC CGG CGC CTG GCC CAG AAT GGC CTG AAA AGT	432
	Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Gly Leu Lys Ser	
	130 135 140	
50	TTC TCC ATT GCC TCT GAC CCA GCC TCC TCA ACC TCC TGC TAC CTG GAA	480
	Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Thr Ser Cys Tyr Leu Glu	
	145 150 155 160	
55	GAG CAT GTG AGC AAG GAG GCT GAG GTC CTG ATA AGC ACG TTG CAG GAG	528
	Glu His Val Ser Lys Glu Ala Glu Val Leu Ile Ser Thr Leu Gln Glu	
	165 170 175	
60	CTG ATG GCA GGG CCT GGG CAC TTT AAC CCC TAC AGG TAT GTG GTG GTA	576
	Leu Met Ala Gly Pro Gly His Phe Asn Pro Tyr Arg Tyr Val Val Val	
	180 185 190	
65	TCA GTG ACC AAT GTC ATC TGT GCC ATT TGC TTT GGC CGG CGC TAT GAC	624
	Ser Val Thr Asn Val Ile Cys Ala Ile Cys Phe Gly Arg Arg Tyr Asp	
	195 200 205	

50

55

	CAC AAC CAC CAA GAA CTG CTT AGC CTA GTC AAC CTG AAT AAT AAT TTC	672
	His Asn His Gln Glu Leu Ser Leu Val Asn Leu Asn Asn Asn Phe	
	210 215 220	
5	GGG GAG GTG GTT GGC TCT GGA AAC CCA GCT GAC TTC ATC CCT ATT CTT	720
	Gly Glu Val Val Gly Ser Gly Asn Pro Ala Asp Phe Ile Pro Ile Leu	
	225 230 235 240	
	CGC TAC CTA CCC AAC CCT TCC CTG AAT GCC TTC AAG GAC CTG AAT GAG	768
	Arg Tyr Leu Pro Asn Pro Ser Leu Asn Ala Phe Lys Asp Leu Asn Glu	
10	245 250 255	
	AAG TTC TAC AGC TTC ATG CAG AAG ATG GTC AAG GAG CAC TAC AAA ACC	816
	Lys Phe Tyr Ser Phe Met Gln Lys Met Val Lys Glu His Tyr Lys Thr	
	260 265 270	
15	TTT GAG AAG GGC CAC ATC CGG GAC ATC ACA GAC AGC CTG ATT GAG CAC	864
	Phe Glu Lys Gly His Ile Arg Asp Ile Thr Asp Ser Leu Ile Glu His	
	275 280 285	
	TGT CAG GAG AAG CAG CTG GAT GAG AAC GCC AAT GTC CAG CTG TCA GAT	912
	Cys Gln Glu Lys Gln Leu Asp Glu Asn Ala Asn Val Gln Leu Ser Asp	
20	290 295 300	
	GAG AAG ATC ATT AAC ATC GTC TTG GAC CTC TTT GGA GCT GGG TTT GAC	960
	Glu Lys Ile Ile Asn Ile Val Leu Asp Leu Phe Gly Ala Gly Phe Asp	
	305 310 315 320	
25	ACA GTC ACA ACT GCT ATC TCC TGG AGC CTC ATG TAT TTG GTG ATG AAC	1008
	Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Met Asn	
	325 330 335	
	CCC AGG GTA CAG AGA AAG ATC CAA GAG GAG CTC GAC ACA GTG ATT GGC	1056
	Pro Arg Val Gln Arg Lys Ile Gln Glu Glu Leu Asp Thr Val Ile Gly	
30	340 345 350	
	AGG TCA CGG CGG CCC CGG CTC TCT GAC AGA TCC CAT CTG CCC TAT ATG	1104
	Arg Ser Arg Arg Pro Arg Leu Ser Asp Arg Ser His Leu Pro Tyr Met	
	355 360 365	
35	GAG GCC TTC ATC CTG GAG ACC TTC CGA CAC TCT TCC TTC GTC CCC TTC	1152
	Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Val Pro Phe	
	370 375 380	
	ACC ATC CCC CAC AGC ACA ACA AGA GAC ACA AGT TTG AAA GGC TTT TAC	1200
	Thr Ile Pro His Ser Thr Thr Arg Asp Thr Ser Leu Lys Gly Phe Tyr	
40	385 390 395 400	
	ATC CCC AAG GGG CGT TGT GTC TTT GTA AAC CAG TGG CAG ATC AAC CAT	1248
	Ile Pro Lys Gly Arg Cys Val Phe Val Asn Gln Trp Gln Ile Asn His	
	405 410 415	
45	GAC CAG AAG CTA TGG GTC AAC CCA TCT GAG TTC CTA CCT GAA CGG TTT	1296
	Asp Gln Lys Leu Trp Val Asn Pro Ser Glu Phe Leu Pro Glu Arg Phe	
	420 425 430	
	CTC ACC CCT GAT GGT GCT ATC GAC AAG GTG TTA AGT GAG AAG GTG ATT	1344
	Leu Thr Pro Asp Gly Ala Ile Asp Lys Val Leu Ser Glu Lys Val Ile	
50	435 440 445	

ATC TTT GGC ATG GGC AAG CGG AAG TGT ATC GGT GAG ACC ATT GCC AGC 1392
 Ile Phe Gly Met Gly Lys Arg Lys Cys Ile Gly Glu Thr Ile Ala Ser
 450 455 460
 5 TGG GAG GTC TTT CTC TTC CTG GCT ATC CTG CTG CAA CGG GTG GAA TTC 1440
 Trp Glu Val Phe Leu Phe Leu Ala Ile Leu Leu Gln Arg Val Glu Phe
 465 470 475 480
 AGC GTG CCA CTG GGC GTG AAG GTG GAC ATG ACC CCC ATC TAT GGG CTA 1488
 Ser Val Pro Leu Gly Val Lys Val Asp Met Thr Pro Ile Tyr Gly Leu
 485 490 495
 10 ACC ATG AAG CAT GCC TGC TGT GAG CAC TTC CAA ATG CAG CTG CGC TCT 1536
 Thr Met Lys His Ala Cys Cys Glu His Phe Gln Met Gln Leu Arg Ser
 500 505 510
 15 TAG 1539

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 512 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

25 Met Leu Phe Pro Ile Ser Met Ser Ala Thr Glu Phe Leu Leu Ala Ser
 1 5 10 15
 Val Ile Phe Cys Leu Val Phe Trp Val Ile Arg Ala Ser Arg Pro Gln
 20 25 30
 30 Val Pro Lys Gly Leu Lys Asn Pro Pro Gly Pro Trp Gly Trp Pro Leu
 35 40 45
 Ile Gly His Met Leu Thr Leu Gly Lys Asn Pro His Leu Ala Leu Ser
 50 55 60
 35 Arg Met Ser Gln Gln Tyr Gly Asp Val Leu Gln Ile Arg Ile Gly Ser
 65 70 75 80
 Thr Pro Val Val Val Leu Ser Gly Leu Asp Thr Ile Arg Gln Ala Leu
 85 90 95
 40 Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr Thr Phe
 100 105 110
 Thr Leu Ile Ser Asn Gly Gln Ser Met Ser Phe Ser Pro Asp Ser Gly
 115 120 125
 45 Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Gly Leu Lys Ser
 130 135 140

50

55

Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Thr Ser Cys Tyr Leu Glu
 145 150 155 160
 5 Glu His Val Ser Lys Glu Ala Glu Val Leu Ile Ser Thr Leu Gln Glu
 165 170 175
 Leu Met Ala Gly Pro Gly His Phe Asn Pro Tyr Arg Tyr Val Val Val
 180 185 190
 10 Ser Val Thr Asn Val Ile Cys Ala Ile Cys Phe Gly Arg Arg Tyr Asp
 195 200 205
 His Asn His Gln Glu Leu Leu Ser Leu Val Asn Leu Asn Asn Asn Phe
 210 215 220
 15 Gly Glu Val Val Gly Ser Gly Asn Pro Ala Asp Phe Ile Pro Ile Leu
 225 230 235 240
 Arg Tyr Leu Pro Asn Pro Ser Leu Asn Ala Phe Lys Asp Leu Asn Glu
 245 250 255
 20 Lys Phe Tyr Ser Phe Met Gln Lys Met Val Lys Glu His Tyr Lys Thr
 260 265 270
 Phe Glu Lys Gly His Ile Arg Asp Ile Thr Asp Ser Leu Ile Glu His
 275 280 285
 25 Cys Gln Glu Lys Gln Leu Asp Glu Asn Ala Asn Val Gln Leu Ser Asp
 290 295 300
 Glu Lys Ile Ile Asn Ile Val Leu Asp Leu Phe Gly Ala Gly Phe Asp
 305 310 315 320
 30 Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Met Asn
 325 330 335
 Pro Arg Val Gln Arg Lys Ile Gln Glu Glu Leu Asp Thr Val Ile Gly
 340 345 350
 35 Arg Ser Arg Arg Pro Arg Leu Ser Asp Arg Ser His Leu Pro Tyr Met
 355 360 365
 Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Val Pro Phe
 370 375 380
 40 Thr Ile Pro His Ser Thr Thr Arg Asp Thr Ser Leu Lys Gly Phe Tyr
 385 390 395 400
 Ile Pro Lys Gly Arg Cys Val Phe Val Asn Gln Trp Gln Ile Asn His
 405 410 415
 45 Asp Gln Lys Leu Trp Val Asn Pro Ser Glu Phe Leu Pro Glu Arg Phe
 420 425 430
 Leu Thr Pro Asp Gly Ala Ile Asp Lys Val Leu Ser Glu Lys Val Ile
 435 440 445
 50 Ile Phe Gly Met Gly Lys Arg Lys Cys Ile Gly Glu Thr Ile Ala Ser
 450 455 460

Trp Glu Val Phe Leu Phe Leu Ala Ile Leu Leu Gln Arg Val Glu Phe
465 470 475 480

5 Ser Val Pro Leu Gly Val Lys Val Asp Met Thr Pro Ile Tyr Gly Leu
485 490 495

Thr Met Lys His Ala Cys Cys Glu His Phe Gln Met Gln Leu Arg Ser
500 505 510

10 (2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1536

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

25	ATG CTT TTC CCA ATC TCC ATG TCG GCC ACG GAG TTT CTT CTG GCC TCT Met Leu Phe Pro Ile Ser Met Ser Ala Thr Glu Phe Leu Leu Ala Ser 1 5 10 15	48
30	GTC ATC TTC TGT CTG GTA TTC TGG GTA ATC AGG GCC TCA AGA CCT CAG Val Ile Phe Cys Leu Val Phe Trp Val Ile Arg Ala Ser Arg Pro Gln 20 25 30	96
35	GTC CCC AAA GGC CTG AAG AAT CCA CCA GGG CCA TGG GGC TGG CCT CTG Val Pro Lys Gly Leu Lys Asn Pro Pro Gly Pro Trp Gly Trp Pro Leu 35 40 45	144
40	ATT GGG CAC ATG CTG ACC CTG GGA AAG AAC CCG CAC CTG GCA CTG TCA Ile Gly His Met Leu Thr Leu Gly Lys Asn Pro His Leu Ala Leu Ser 50 55 60	192
45	AGG ATG AGC CAG CAG TAT GGG GAC GTG CTG CAG ATC CGA ATT GGC TCC Arg Met Ser Gln Gln Tyr Gly Asp Val Leu Gln Ile Arg Ile Gly Ser 65 70 75 80	240
50	ACA CCC GTG GTG GTG CTG AGC GGC CTG GAC ACC ATC CGG CAG GCC CTG Thr Pro Val Val Val Leu Ser Gly Leu Asp Thr Ile Arg Gln Ala Leu 85 90 95	288
55	GTG CGG CAG GGC GAT GAT TTC AAG GGC CGG CCC GAC CTC TAC ACC TTC Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr Thr Phe 100 105 110	336
60	ACC CTC ATC AGT AAT GGT CAG AGC ATG TCC TTC AGC CCA GAC TCT GGA Thr Leu Ile Ser Asn Gly Gln Ser Met Ser Phe Ser Pro Asp Ser Gly 115 120 125	384

50

55

	CCA GTG TGG GCT GCC CGC CGG CGC CTG GCC CAG AAT GGC CTG AAA AGT	432
	Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Gly Leu Lys Ser	
	130 135 140	
5	TTC TCC ATT GCC TCT GAC CCA GCC TCC TCA ACC TCC TGC TAC CTG GAA	480
	Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Thr Ser Cys Tyr Leu Glu	
	145 150 155 160	
	GAG CAT GTG AGC AAG GAG GCT GAG GTC CTG ATA AGC ACG TTG CAG GAG	528
10	Glu His Val Ser Lys Glu Ala Glu Val Leu Ile Ser Thr Leu Gln Glu	
	165 170 175	
	CTG ATG GCA GGG CCT GGG CAC TTT AAC CCC TAC AGG TAT GTG GTG GTA	576
	Leu Met Ala Gly Pro Gly His Phe Asn Pro Tyr Arg Tyr Val Val Val	
	180 185 190	
15	TCA GTG ACC AAT GTC ATC TGT GCC ATT TGC TTT GGC CGG CGC TAT GAC	624
	Ser Val Thr Asn Val Ile Cys Ala Ile Cys Phe Gly Arg Arg Tyr Asp	
	195 200 205	
	CAC AAC CAC CAA GAA CTG CTT AGC CTA GTC AAC CTG AAT AAT AAT TTC	672
20	His Asn His Gln Glu Leu Leu Ser Leu Val Asn Leu Asn Asn Asn Phe	
	210 215 220	
	GGG GAG GTG GTT GGC TCT GGA AAC CCA GCT GAC TTC ATC CCT ATT CTT	720
	Gly Glu Val Val Gly Ser Gly Asn Pro Ala Asp Phe Ile Pro Ile Leu	
	225 230 235 240	
25	CGC TAC CTA CCC AAC CCT TCC CTG AAT GCC TTC AAG GAC CTG AAT GAG	768
	Arg Tyr Leu Pro Asn Pro Ser Leu Asn Ala Phe Lys Asp Leu Asn Glu	
	245 250 255	
	AAG TTC TAC AGC TTC ATG CAG AAG ATG GTC AAG GAG CAC TAC AAA ACC	816
30	Lys Phe Tyr Ser Phe Met Gln Lys Met Val Lys Glu His Tyr Lys Thr	
	260 265 270	
	TTT GAG AAG GGC CAC ATC CGG GAC ATC ACA GAC AGC CTG ATT GAG CAC	864
	Phe Glu Lys Gly His Ile Arg Asp Ile Thr Asp Ser Leu Ile Glu His	
	275 280 285	
35	TGT CAG GAG AAG CAG CTG GAT GAG AAC GCC AAT GTC CAG CTG TCA GAT	912
	Cys Gln Glu Lys Gln Leu Asp Glu Asn Ala Asn Val Gln Leu Ser Asp	
	290 295 300	
	GAG AAG ATC ATT AAC ATC GTC TTG GAC CTC TTT GGA GCT GGG TTT GAC	960
40	Glu Lys Ile Ile Asn Ile Val Leu Asp Leu Phe Gly Ala Gly Phe Asp	
	305 310 315 320	
	ACA GTC ACA ACT GCT ATC TCC TGG AGC CTC ATG TAT TTG GTG ATG AAC	1008
	Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Met Asn	
	325 330 335	
45	CCC AGG GTA CAG AGA AAG ATC CAA GAG GAG CTC GAC ACA GTG ATT GGC	1056
	Pro Arg Val Gln Arg Lys Ile Gln Glu Glu Leu Asp Thr Val Ile Gly	
	340 345 350	
	AGG TCA CGG CGG CCC CGG CTC TCT GAC AGA TCC CAT CTG CCC TAT ATG	1104
50	Arg Ser Arg Arg Pro Arg Leu Ser Asp Arg Ser His Leu Pro Tyr Met	
	355 360 365	

	GAG GCC TTC ATC CTG GAG ACC TTC CGA CAC TCT TCC TTC GTC CCC TTC	1152
	Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Val Pro Phe	
	370 375 380	
5	ACC ATC CCC CAC AGC ACA ACA AGA GAC ACA AGT TTG AAA GGC TTT TAC	1200
	Thr Ile Pro His Ser Thr Thr Arg Asp Thr Ser Leu Lys Gly Phe Tyr	
	385 390 395 400	
	ATC CCC AAG GGG CGT TGT GTC TTT GTA AAC CAG TGG CAG ATC AAC CAT	1248
10	Ile Pro Lys Gly Arg Cys Val Phe Val Asn Gln Trp Gln Ile Asn His	
	405 410 415	
	GAC CAG AAG CTA TGG GTC AAC CCA TCT GAG TTC CTA CCT GAA CGG TTT	1296
	Asp Gln Lys Leu Trp Val Asn Pro Ser Glu Phe Leu Pro Glu Arg Phe	
	420 425 430	
15	CTC ACC CCT GAT GGT GCT ATC GAC AAG GTG TTA AGT GAG AAG GTG ATT	1344
	Leu Thr Pro Asp Gly Ala Ile Asp Lys Val Leu Ser Glu Lys Val Ile	
	435 440 445	
	ATC TTT GGC ATG GGC AAG CGG AAG TGT ATC GGT GAG ACC ATT GCC CGC	1392
20	Ile Phe Gly Met Gly Lys Arg Lys Cys Ile Gly Glu Thr Ile Ala Arg	
	450 455 460	
	TGG GAG GTC TTT CTC TTC CTG GCT ATC CTG CTG CAA CGG GTG GAA TTC	1440
	Trp Glu Val Phe Leu Phe Leu Ala Ile Leu Leu Gln Arg Val Glu Phe	
	465 470 475 480	
25	AGC GTG CCA CTG GGC GTG AAG GTG GAC ATG ACC CCC ATC TAT GGG CTA	1488
	Ser Val Pro Leu Gly Val Lys Val Asp Met Thr Pro Ile Tyr Gly Leu	
	485 490 495	
	ACC ATG AAG CAT GCC TGC TGT GAG CAC TTC CAA ATG CAG CTG CGC TCT	1536
30	Thr Met Lys His Ala Cys Cys Glu His Phe Gln Met Gln Leu Arg Ser	
	500 505 510	
	TAG	1539

(2) INFORMATION FOR SEQ ID NO: 12:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 512 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met	Leu	Phe	Pro	Ile	Ser	Met	Ser	Ala	Thr	Glu	Phe	Leu	Leu	Ala	Ser
1				5					10					15	
Val	Ile	Phe	Cys	Leu	Val	Phe	Trp	Val	Ile	Arg	Ala	Ser	Arg	Pro	Gln
			20					25					30		
Val	Pro	Lys	Gly	Leu	Lys	Asn	Pro	Pro	Gly	Pro	Trp	Gly	Trp	Pro	Leu
		35					40					45			

50

55

Ile Gly His Met Leu Thr Leu Gly Lys Asn Pro His Leu Ala Leu Ser
 50 55 60
 5 Arg Met Ser Gln Gln Tyr Gly Asp Val Leu Gln Ile Arg Ile Gly Ser
 65 70 75 80
 Thr Pro Val Val Val Leu Ser Gly Leu Asp Thr Ile Arg Gln Ala Leu
 85 90 95
 10 Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr Thr Phe
 100 105 110
 Thr Leu Ile Ser Asn Gly Gln Ser Met Ser Phe Ser Pro Asp Ser Gly
 115 120 125
 15 Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Gly Leu Lys Ser
 130 135 140
 Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Thr Ser Cys Tyr Leu Glu
 145 150 155 160
 20 Glu His Val Ser Lys Glu Ala Glu Val Leu Ile Ser Thr Leu Gln Glu
 165 170 175
 Leu Met Ala Gly Pro Gly His Phe Asn Pro Tyr Arg Tyr Val Val Val
 180 185 190
 25 Ser Val Thr Asn Val Ile Cys Ala Ile Cys Phe Gly Arg Arg Tyr Asp
 195 200 205
 His Asn His Gln Glu Leu Leu Ser Leu Val Asn Leu Asn Asn Asn Phe
 210 215 220
 30 Gly Glu Val Val Gly Ser Gly Asn Pro Ala Asp Phe Ile Pro Ile Leu
 225 230 235 240
 Arg Tyr Leu Pro Asn Pro Ser Leu Asn Ala Phe Lys Asp Leu Asn Glu
 245 250 255
 35 Lys Phe Tyr Ser Phe Met Gln Lys Met Val Lys Glu His Tyr Lys Thr
 260 265 270
 40 Phe Glu Lys Gly His Ile Arg Asp Ile Thr Asp Ser Leu Ile Glu His
 275 280 285
 Cys Gln Glu Lys Gln Leu Asp Glu Asn Ala Asn Val Gln Leu Ser Asp
 290 295 300
 45 Glu Lys Ile Ile Asn Ile Val Leu Asp Leu Phe Gly Ala Gly Phe Asp
 305 310 315 320
 Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Met Asn
 325 330 335
 50 Pro Arg Val Gln Arg Lys Ile Gln Glu Glu Leu Asp Thr Val Ile Gly
 340 345 350

Arg Ser Arg Arg Pro Arg Leu Ser Asp Arg Ser His Leu Pro Tyr Met
 355 360 365
 5 Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Val Pro Phe
 370 375 380
 Thr Ile Pro His Ser Thr Thr Arg Asp Thr Ser Leu Lys Gly Phe Tyr
 385 390 395 400
 10 Ile Pro Lys Gly Arg Cys Val Phe Val Asn Gln Trp Gln Ile Asn His
 405 410 415
 Asp Gln Lys Leu Trp Val Asn Pro Ser Glu Phe Leu Pro Glu Arg Phe
 420 425 430
 15 Leu Thr Pro Asp Gly Ala Ile Asp Lys Val Leu Ser Glu Lys Val Ile
 435 440 445
 Ile Phe Gly Met Gly Lys Arg Lys Cys Ile Gly Glu Thr Ile Ala Arg
 450 455 460
 20 Trp Glu Val Phe Leu Phe Leu Ala Ile Leu Leu Gln Arg Val Glu Phe
 465 470 475 480
 Ser Val Pro Leu Gly Val Lys Val Asp Met Thr Pro Ile Tyr Gly Leu
 485 490 495
 25 Thr Met Lys His Ala Cys Cys Glu His Phe Gln Met Gln Leu Arg Ser
 500 505 510

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 1539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- 35 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1536

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

40 ATG CTT TTC CCA ATC TCC ATG TCG GCC ACG GAG TTT CTT CTG GCC TCT 48
 Met Leu Phe Pro Ile Ser Met Ser Ala Thr Glu Phe Leu Leu Ala Ser
 1 5 10 15
 GTC ATC TTC TGT CTG GTA TTC TGG GTA ATC AGG GCC TCA AGA CCT CAG 96
 Val Ile Phe Cys Leu Val Phe Trp Val Ile Arg Ala Ser Arg Pro Gln
 45 20 25 30
 GTC CCC AAA GGC CTG AAG AAT CCA CCA GGG CCA TGG GGC TGG CCT CTG 144
 Val Pro Lys Gly Leu Lys Asn Pro Pro Gly Pro Trp Gly Trp Pro Leu
 35 40 45

50

55

	ATT GGG CAC ATG CTG ACC CTG GGA AAG AAC CCG CAC CTG GCA CTG TCA	192
	Ile Gly His Met Leu Thr Leu Gly Lys Asn Pro His Leu Ala Leu Ser	
	50 55 60	
5	AGG ATG AGC CAG CAG TAT GGG GAC GTG CTG CAG ATC CGA ATT GGC TCC	240
	Arg Met Ser Gln Gln Tyr Gly Asp Val Leu Gln Ile Arg Ile Gly Ser	
	65 70 75 80	
	ACA CCC GTG GTG GTG CTG AGC GGC CTG GAC ACC ATC CGG CAG GCC CTG	288
	Thr Pro Val Val Leu Ser Gly Leu Asp Thr Ile Arg Gln Ala Leu	
	85 90 95	
10	GTG CGG CAG GGC GAT GAT TTC AAG GGC CGG CCC GAC CTC TAC ACC TTC	336
	Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr Thr Phe	
	100 105 110	
15	ACC CTC ATC AGT AAT GGT CAG AGC ATG TCC TTC AGC CCA GAC TCT GGA	384
	Thr Leu Ile Ser Asn Gly Gln Ser Met Ser Phe Ser Pro Asp Ser Gly	
	115 120 125	
	CCA GTG TGG GCT GCC CGC CGG CGC CTG GCC CAG AAT GGC CTG AAA AGT	432
	Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Gly Leu Lys Ser	
	130 135 140	
20	TTC TCC ATT GCC TCT GAC CCA GCC TCC TCA ACC TCC TGC TAC CTG GAA	480
	Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Thr Ser Cys Tyr Leu Glu	
	145 150 155 160	
25	GAG CAT GTG AGC AAG GAG GCT GAG GTC CTG ATA AGC ACG TTG CAG GAG	528
	Glu His Val Ser Lys Glu Ala Glu Val Leu Ile Ser Thr Leu Gln Glu	
	165 170 175	
	CTG ATG GCA GGG CCT GGG CAC TTT AAC CCC TAC AGG TAT GTG GTG GTA	576
	Leu Met Ala Gly Pro Gly His Phe Asn Pro Tyr Arg Tyr Val Val Val	
	180 185 190	
30	TCA GTG ACC AAT GTC ATC TGT GCC ATT TGC TTT GGC CGG CGC TAT GAC	624
	Ser Val Thr Asn Val Ile Cys Ala Ile Cys Phe Gly Arg Arg Tyr Asp	
	195 200 205	
35	CAC AAC CAC CAA GAA CTG CTT AGC CTA GTC AAC CTG AAT AAT AAT TTC	672
	His Asn His Gln Glu Leu Leu Ser Leu Val Asn Leu Asn Asn Asn Phe	
	210 215 220	
	GGG GAG GTG GTT GGC TCT GGA AAC CCA GCT GAC TTC ATC CCT ATT CTT	720
	Gly Glu Val Val Gly Ser Gly Asn Pro Ala Asp Phe Ile Pro Ile Leu	
	225 230 235 240	
40	CGC TAC CTA CCC AAC CCT TCC CTG AAT GCC TTC AAG GAC CTG AAT GAG	768
	Arg Tyr Leu Pro Asn Pro Ser Leu Asn Ala Phe Lys Asp Leu Asn Glu	
	245 250 255	
45	AAG TTC TAC AGC TTC ATG CAG AAG ATG GTC AAG GAG CAC TAC AAA ACC	816
	Lys Phe Tyr Ser Phe Met Gln Lys Met Val Lys Glu His Tyr Lys Thr	
	260 265 270	
	TTT GAG AAG GGC CAC ATC CGG GAC ATC ACA GAC AGC CTG ATT GAG CAC	864
	Phe Glu Lys Gly His Ile Arg Asp Ile Thr Asp Ser Leu Ile Glu His	
	275 280 285	
50		
55		

	TGT CAG GAG AAG CAG CTG GAT GAG AAC GCC AAT GTC CAG CTG TCA GAT	912
	Cys Gln Glu Lys Gln Leu Asp Glu Asn Ala Asn Val Gln Leu Ser Asp	
	290 295 300	
5	GAG AAG ATC ATT AAC ATC GTC TTG GAC CTC TTT GGA GCT GGG TTT GAC	960
	Glu Lys Ile Ile Asn Ile Val Leu Asp Leu Phe Gly Ala Gly Phe Asp	
	305 310 315 320	
	ACA GTC ACA ACT GCT ATC TCC TGG AGC CTC ATG TAT TTG GTG ATG AAC	1008
10	Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Met Asn	
	325 330 335	
	CCC AGG GTA CAG AGA AAG ATC CAA GAG GAG CTC GAC ACA GTG ATT GGC	1056
	Pro Arg Val Gln Arg Lys Ile Gln Glu Glu Leu Asp Thr Val Ile Gly	
	340 345 350	
15	AGG TCA CGG CGG CCC CGG CTC TCT GAC AGA TCC CAT CTG CCC TAT ATG	1104
	Arg Ser Arg Arg Pro Arg Leu Ser Asp Arg Ser His Leu Pro Tyr Met	
	355 360 365	
	GAG GCC TTC ATC CTG GAG ACC TTC CGA CAC TCT TCC TTC GTC CCC TTC	1152
20	Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Val Pro Phe	
	370 375 380	
	ACC ATC CCC CAC AGC ACA ACA AGA GAC ACA AGT TTG AAA GGC TTT TAC	1200
	Thr Ile Pro His Ser Thr Thr Arg Asp Thr Ser Leu Lys Gly Phe Tyr	
	385 390 395 400	
25	ATC CCC AAG GGG CGT TGT GTC TTT GTA AAC CAG TGG CAG ATC AAC CAT	1248
	Ile Pro Lys Gly Arg Cys Val Phe Val Asn Gln Trp Gln Ile Asn His	
	405 410 415	
	GAC CAG AAG CTA TGG GTC AAC CCA TCT GAG TTC CTA CCT GAA CGG TTT	1296
30	Asp Gln Lys Leu Trp Val Asn Pro Ser Glu Phe Leu Pro Glu Arg Phe	
	420 425 430	
	CTC ACC CCT GAT GGT GCT ATC GAC AAG GTG TTA AGT GAG AAG GTG ATT	1344
	Leu Thr Pro Asp Gly Ala Ile Asp Lys Val Leu Ser Glu Lys Val Ile	
	435 440 445	
35	ATC TTT GGC ATG GGC AAG CGG AAG TGT ATC GGT GAG ACC GTT GCC CGC	1392
	Ile Phe Gly Met Gly Lys Arg Lys Cys Ile Gly Glu Thr Val Ala Arg	
	450 455 460	
	TGG GAG GTC TTT CTC TTC CTG GCT ATC CTG CTG CAA CGG GTG GAA TTC	1440
40	Trp Glu Val Phe Leu Phe Leu Ala Ile Leu Leu Gln Arg Val Glu Phe	
	465 470 475 480	
	AGC GTG CCA CTG GGC GTG AAG GTG GAC ATG ACC CCC ATC TAT GGG CTA	1488
	Ser Val Pro Leu Gly Val Lys Val Asp Met Thr Pro Ile Tyr Gly Leu	
	485 490 495	
45	ACC ATG AAG CAT GCC TGC TGT GAG CAC TTC CAA ATG CAG CTG CGC TCT	1536
	Thr Met Lys His Ala Cys Cys Glu His Phe Gln Met Gln Leu Arg Ser	
	500 505 510	
	TAG	1539

50

55

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 512 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

10 Met Leu Phe Pro Ile Ser Met Ser Ala Thr Glu Phe Leu Leu Ala Ser
 1 5 10 15
 Val Ile Phe Cys Leu Val Phe Trp Val Ile Arg Ala Ser Arg Pro Gln
 20 25 30
 15 Val Pro Lys Gly Leu Lys Asn Pro Pro Gly Pro Trp Gly Trp Pro Leu
 35 40 45
 Ile Gly His Met Leu Thr Leu Gly Lys Asn Pro His Leu Ala Leu Ser
 50 55 60
 20 Arg Met Ser Gln Gln Tyr Gly Asp Val Leu Gln Ile Arg Ile Gly Ser
 65 70 75 80
 Thr Pro Val Val Val Leu Ser Gly Leu Asp Thr Ile Arg Gln Ala Leu
 85 90 95
 25 Val Arg Gln Gly Asp Asp Phe Lys Gly Arg Pro Asp Leu Tyr Thr Phe
 100 105 110
 Thr Leu Ile Ser Asn Gly Gln Ser Met Ser Phe Ser Pro Asp Ser Gly
 115 120 125
 30 Pro Val Trp Ala Ala Arg Arg Arg Leu Ala Gln Asn Gly Leu Lys Ser
 130 135 140
 Phe Ser Ile Ala Ser Asp Pro Ala Ser Ser Thr Ser Cys Tyr Leu Glu
 145 150 155 160
 35 Glu His Val Ser Lys Glu Ala Glu Val Leu Ile Ser Thr Leu Gln Glu
 165 170 175
 Leu Met Ala Gly Pro Gly His Phe Asn Pro Tyr Arg Tyr Val Val Val
 180 185 190
 40 Ser Val Thr Asn Val Ile Cys Ala Ile Cys Phe Gly Arg Arg Tyr Asp
 195 200 205
 His Asn His Gln Glu Leu Leu Ser Leu Val Asn Leu Asn Asn Asn Phe
 210 215 220
 45 Gly Glu Val Val Gly Ser Gly Asn Pro Ala Asp Phe Ile Pro Ile Leu
 225 230 235 240
 Arg Tyr Leu Pro Asn Pro Ser Leu Asn Ala Phe Lys Asp Leu Asn Glu
 245 250 255

Lys Phe Tyr Ser Phe Met Gln Lys Met Val Lys Glu His Tyr Lys Thr
 260 265 270
 5 Phe Glu Lys Gly His Ile Arg Asp Ile Thr Asp Ser Leu Ile Glu His
 275 280 285
 Cys Gln Glu Lys Gln Leu Asp Glu Asn Ala Asn Val Gln Leu Ser Asp
 290 295 300
 10 Glu Lys Ile Ile Asn Ile Val Leu Asp Leu Phe Gly Ala Gly Phe Asp
 305 310 315 320
 Thr Val Thr Thr Ala Ile Ser Trp Ser Leu Met Tyr Leu Val Met Asn
 325 330 335
 15 Pro Arg Val Gln Arg Lys Ile Gln Glu Glu Leu Asp Thr Val Ile Gly
 340 345 350
 Arg Ser Arg Arg Pro Arg Leu Ser Asp Arg Ser His Leu Pro Tyr Met
 355 360 365
 20 Glu Ala Phe Ile Leu Glu Thr Phe Arg His Ser Ser Phe Val Pro Phe
 370 375 380
 Thr Ile Pro His Ser Thr Thr Arg Asp Thr Ser Leu Lys Gly Phe Tyr
 385 390 395 400
 25 Ile Pro Lys Gly Arg Cys Val Phe Val Asn Gln Trp Gln Ile Asn His
 405 410 415
 Asp Gln Lys Leu Trp Val Asn Pro Ser Glu Phe Leu Pro Glu Arg Phe
 420 425 430
 30 Leu Thr Pro Asp Gly Ala Ile Asp Lys Val Leu Ser Glu Lys Val Ile
 435 440 445
 Ile Phe Gly Met Gly Lys Arg Lys Cys Ile Gly Glu Thr Val Ala Arg
 450 455 460
 35 Trp Glu Val Phe Leu Phe Leu Ala Ile Leu Leu Gln Arg Val Glu Phe
 465 470 475 480
 Ser Val Pro Leu Gly Val Lys Val Asp Met Thr Pro Ile Tyr Gly Leu
 485 490 495
 40 Thr Met Lys His Ala Cys Cys Glu His Phe Gln Met Gln Leu Arg Ser
 500 505 510

45 (2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1485 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

55

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1482

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

ATG CTG GCC TCA GGG ATG CTT CTG GTG GCC TTG CTG GTC TGC CTG ACT	48
Met Leu Ala Ser Gly Met Leu Leu Val Ala Leu Leu Val Cys Leu Thr	
1 5 10 15	
GTG ATG GTC TTG ATG TCT GTT TGG CAG CAG AGG AAG AGC AAG GGG AAG	96
Val Met Val Leu Met Ser Val Trp Gln Gln Arg Lys Ser Lys Gly Lys	
20 25 30	
CTG CCT CCG GGA CCC ACC CCA TTG CCC TTC ATT GGA AAC TAC CTG CAG	144
Leu Pro Pro Gly Pro Thr Pro Leu Pro Phe Ile Gly Asn Tyr Leu Gln	
35 40 45	
CTG AAC ACA GAG CAG ATG TAC AAC TCC CTC ATG AAG ATC AGT GAG CGC	192
Leu Asn Thr Glu Gln Met Tyr Asn Ser Leu Met Lys Ile Ser Glu Arg	
50 55 60	
TAT GGC CCC GTG TTC ACC ATT CAC TTG GGG CCC CGG CGG GTC GTG GTG	240
Tyr Gly Pro Val Phe Thr Ile His Leu Gly Pro Arg Arg Val Val Val	
65 70 75 80	
CTG TGT GGA CAT GAT GCC GTC AGG GAG GCT CTG GTG GAC CAG GCT GAG	288
Leu Cys Gly His Asp Ala Val Arg Glu Ala Leu Val Asp Gln Ala Glu	
85 90 95	
GAG TTC AGC GGG CGA GGC GAG CAA GCC ACC TTC GAC TGG GTC TTC AAA	336
Glu Phe Ser Gly Arg Gly Glu Gln Ala Thr Phe Asp Trp Val Phe Lys	
100 105 110	
GGC TAT GGC GTG GTA TTC AGC AAC GGG GAG CGC GCC AAG CAG CTC CGG	384
Gly Tyr Gly Val Val Phe Ser Asn Gly Glu Arg Ala Lys Gln Leu Arg	
115 120 125	
CGC TTC TCC ATC GCC ACC CTG CGG GAC TTC GGG GTG GGC AAG CGA GGC	432
Arg Phe Ser Ile Ala Thr Leu Arg Asp Phe Gly Val Gly Lys Arg Gly	
130 135 140	
ATC GAG GAG CGC ATC CAG GAG GAG GCG GGC TTC CTC ATC GAC GCC CTC	480
Ile Glu Glu Arg Ile Gln Glu Glu Ala Gly Phe Leu Ile Asp Ala Leu	
145 150 155 160	
CGG GGC ACT GGC GGC GCC AAT ATC GAT CCC ACC TTC TTC CTG AGC CGC	528
Arg Gly Thr Gly Gly Ala Asn Ile Asp Pro Thr Phe Phe Leu Ser Arg	
165 170 175	
ACA GTC TCC AAT GTC ATC AGC TCC ATT GTC TTT GGG GAC CGC TTT GAC	576
Thr Val Ser Asn Val Ile Ser Ser Ile Val Phe Gly Asp Arg Phe Asp	
180 185 190	
TAT AAG GAC AAA GAG TTC CTG TCA CTG TTG CGC ATG ATG CTA GGA ATC	624
Tyr Lys Asp Lys Glu Phe Leu Ser Leu Leu Arg Met Met Leu Gly Ile	
195 200 205	

50

55

	TTC	CAG	TTC	ACG	TCA	ACC	TCC	ACG	GGG	CAG	CTC	TAT	GAG	ATG	TTC	TCT	672
	Phe	Gln	Phe	Thr	Ser	Thr	Ser	Thr	Gly	Gln	Leu	Tyr	Glu	Met	Phe	Ser	
	210						215				220						
5	TCG	GTG	ATG	AAA	CAC	CTG	CCA	GGA	CCA	CAG	CAA	CAG	GCC	TTT	CAG	TTG	720
	Ser	Val	Met	Lys	His	Leu	Pro	Gly	Pro	Gln	Gln	Gln	Ala	Phe	Gln	Leu	
	225					230				235						240	
	CTG	CAA	GGG	CTG	GAG	GAC	TTC	ATA	GCC	AAG	AAG	GTG	GAG	CAC	AAC	CAG	768
	Leu	Gln	Gly	Leu	Glu	Asp	Phe	Ile	Ala	Lys	Lys	Val	Glu	His	Asn	Gln	
					245					250					255		
10	CGC	ACG	CTG	GAT	CCC	AAT	TCC	CCA	CGG	GAC	TTC	ATT	GAC	TCC	TTT	CTC	816
	Arg	Thr	Leu	Asp	Pro	Asn	Ser	Pro	Arg	Asp	Phe	Ile	Asp	Ser	Phe	Leu	
				260					265					270			
	ATC	CGC	ATG	CAG	GAG	GAG	GAG	AAG	AAC	CCC	AAC	ACG	GAG	TTC	TAC	TTG	864
15	Ile	Arg	Met	Gln	Glu	Glu	Glu	Lys	Asn	Pro	Asn	Thr	Glu	Phe	Tyr	Leu	
			275					280					285				
	AAA	AAC	CTG	GTG	ATG	ACC	ACG	TTG	AAC	CTC	TTC	ATT	GGG	GGC	ACC	GAG	912
	Lys	Asn	Leu	Val	Met	Thr	Thr	Leu	Asn	Leu	Phe	Ile	Gly	Gly	Thr	Glu	
		290					295					300					
20	ACC	GTC	AGC	ACC	ACC	CTG	CGC	TAT	GGC	TTC	TTG	CTG	CTC	ATG	AAG	CAC	960
	Thr	Val	Ser	Thr	Thr	Leu	Arg	Tyr	Gly	Phe	Leu	Leu	Leu	Met	Lys	His	
	305					310					315					320	
	CCA	GAG	GTG	GAG	GCC	AAG	GTC	CAT	GAG	GAG	ATT	GAC	AGA	GTG	ATC	GGC	1008
25	Pro	Glu	Val	Glu	Ala	Lys	Val	His	Glu	Glu	Ile	Asp	Arg	Val	Ile	Gly	
					325					330					335		
	AAG	AAC	CGG	CAG	CCC	AAG	TTT	GAG	GAC	CGG	GCC	AAG	ATG	CCC	TAC	ATG	1056
	Lys	Asn	Arg	Gln	Pro	Lys	Phe	Glu	Asp	Arg	Ala	Lys	Met	Pro	Tyr	Met	
				340					345					350			
30	GAG	GCA	GTG	ATC	CAC	GAG	ATC	CAA	AGA	TTT	GGA	GAC	GTG	ATC	CCC	ATG	1104
	Glu	Ala	Val	Ile	His	Glu	Ile	Gln	Arg	Phe	Gly	Asp	Val	Ile	Pro	Met	
			355					360					365				
	AGT	TTG	GCC	CGC	AGA	GTC	AAA	AAG	GAC	ACC	AAG	TTT	CGG	GAT	TTC	TTC	1152
35	Ser	Leu	Ala	Arg	Arg	Val	Lys	Lys	Asp	Thr	Lys	Phe	Arg	Asp	Phe	Phe	
		370					375					380					
	CTC	CCT	AAG	GGC	ACC	GAA	GTG	TAC	CCT	ATG	CTG	GGC	TCT	GTG	CTG	AGA	1200
	Leu	Pro	Lys	Gly	Thr	Glu	Val	Tyr	Pro	Met	Leu	Gly	Ser	Val	Leu	Arg	
	385					390					395					400	
40	GAC	CCC	AGT	TTC	TTC	TCC	AAC	CCC	CAG	GAC	TTC	AAT	CCC	CAG	CAC	TTC	1248
	Asp	Pro	Ser	Phe	Phe	Ser	Asn	Pro	Gln	Asp	Phe	Asn	Pro	Gln	His	Phe	
					405					410					415		
	CTG	AAT	GAG	AAG	GGG	CAG	TTT	AAG	AAG	AGT	GAT	GCT	TTT	GTG	CCC	TTT	1296
45	Leu	Asn	Glu	Lys	Gly	Gln	Phe	Lys	Lys	Ser	Asp	Ala	Phe	Val	Pro	Phe	
				420					425					430			
	TCC	ATC	GGA	AAG	CGG	AAC	TGT	TTC	GGA	GAA	GGC	CTG	GCC	AGA	ATG	GAG	1344
	Ser	Ile	Gly	Lys	Arg	Asn	Cys	Phe	Gly	Glu	Gly	Leu	Ala	Arg	Met	Glu	
			435				440						445				

50

55

CTC TTT CTC TTC TTC ACC ACC GTC ATG CAG AAC TTC CGC CTC AAG TCC 1392
 Leu Phe Leu Phe Phe Thr Thr Val Met Gln Asn Phe Arg Leu Lys Ser
 450 455 460

5 TCC CAG TCA CCT AAG GAC ATT GAC GTG TCC CCC AGA CAC GTG GGC TTT 1440
 Ser Gln Ser Pro Lys Asp Ile Asp Val Ser Pro Arg His Val Gly Phe
 465 470 475 480

GCC ACG ATC CCA CGA AAC TAC ACC ATG AGC TTC CTG CCC CGC 1482
 Ala Thr Ile Pro Arg Asn Tyr Thr Met Ser Phe Leu Pro Arg
 10 485 490

TGA 1485

(2) INFORMATION FOR SEQ ID NO: 16:

- 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 494 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
- 20 (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Leu Ala Ser Gly Met Leu Leu Val Ala Leu Leu Val Cys Leu Thr
 1 5 10 15

25 Val Met Val Leu Met Ser Val Trp Gln Gln Arg Lys Ser Lys Gly Lys
 20 25 30

Leu Pro Pro Gly Pro Thr Pro Leu Pro Phe Ile Gly Asn Tyr Leu Gln
 35 40 45

30 Leu Asn Thr Glu Gln Met Tyr Asn Ser Leu Met Lys Ile Ser Glu Arg
 50 55 60

Tyr Gly Pro Val Phe Thr Ile His Leu Gly Pro Arg Arg Val Val Val
 65 70 75 80

35 Leu Cys Gly His Asp Ala Val Arg Glu Ala Leu Val Asp Gln Ala Glu
 85 90 95

Glu Phe Ser Gly Arg Gly Glu Gln Ala Thr Phe Asp Trp Val Phe Lys
 100 105 110

40 Gly Tyr Gly Val Val Phe Ser Asn Gly Glu Arg Ala Lys Gln Leu Arg
 115 120 125

Arg Phe Ser Ile Ala Thr Leu Arg Asp Phe Gly Val Gly Lys Arg Gly
 130 135 140

45 Ile Glu Glu Arg Ile Gln Glu Glu Ala Gly Phe Leu Ile Asp Ala Leu
 145 150 155 160

Arg Gly Thr Gly Gly Ala Asn Ile Asp Pro Thr Phe Phe Leu Ser Arg
 165 170 175

50

55

Thr Val Ser Asn Val Ile Ser Ser Ile Val Phe Gly Asp Arg Phe Asp
 180 185 190
 5 Tyr Lys Asp Lys Glu Phe Leu Ser Leu Leu Arg Met Met Leu Gly Ile
 195 200 205
 Phe Gln Phe Thr Ser Thr Ser Thr Gly Gln Leu Tyr Glu Met Phe Ser
 210 215 220
 10 Ser Val Met Lys His Leu Pro Gly Pro Gln Gln Gln Ala Phe Gln Leu
 225 230 235 240
 Leu Gln Gly Leu Glu Asp Phe Ile Ala Lys Lys Val Glu His Asn Gln
 245 250 255
 15 Arg Thr Leu Asp Pro Asn Ser Pro Arg Asp Phe Ile Asp Ser Phe Leu
 260 265 270
 Ile Arg Met Gln Glu Glu Glu Lys Asn Pro Asn Thr Glu Phe Tyr Leu
 275 280 285
 20 Lys Asn Leu Val Met Thr Thr Leu Asn Leu Phe Ile Gly Gly Thr Glu
 290 295 300
 Thr Val Ser Thr Thr Leu Arg Tyr Gly Phe Leu Leu Leu Met Lys His
 305 310 315 320
 25 Pro Glu Val Glu Ala Lys Val His Glu Glu Ile Asp Arg Val Ile Gly
 325 330 335
 Lys Asn Arg Gln Pro Lys Phe Glu Asp Arg Ala Lys Met Pro Tyr Met
 340 345 350
 30 Glu Ala Val Ile His Glu Ile Gln Arg Phe Gly Asp Val Ile Pro Met
 355 360 365
 Ser Leu Ala Arg Arg Val Lys Lys Asp Thr Lys Phe Arg Asp Phe Phe
 370 375 380
 35 Leu Pro Lys Gly Thr Glu Val Tyr Pro Met Leu Gly Ser Val Leu Arg
 385 390 395 400
 Asp Pro Ser Phe Phe Ser Asn Pro Gln Asp Phe Asn Pro Gln His Phe
 405 410 415
 40 Leu Asn Glu Lys Gly Gln Phe Lys Lys Ser Asp Ala Phe Val Pro Phe
 420 425 430
 Ser Ile Gly Lys Arg Asn Cys Phe Gly Glu Gly Leu Ala Arg Met Glu
 435 440 445
 45 Leu Phe Leu Phe Phe Thr Thr Val Met Gln Asn Phe Arg Leu Lys Ser
 450 455 460
 50 Ser Gln Ser Pro Lys Asp Ile Asp Val Ser Pro Arg His Val Gly Phe
 465 470 475 480

Ala Thr Ile Pro Arg Asn Tyr Thr Met Ser Phe Leu Pro Arg
 485 490

(2) INFORMATION FOR SEQ ID NO: 17:

5

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1485 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

10

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1482

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATG CTG GCC TCA GGG ATG CTT CTG GTG GCC TTG CTG GTC TGC CTG ACT	48
Met Leu Ala Ser Gly Met Leu Leu Val Ala Leu Leu Val Cys Leu Thr	
1 5 10 15	
GTG ATG GTC TTG ATG TCT GTT TGG CAG CAG AGG AAG AGC AAG GGG AAG	96
Val Met Val Leu Met Ser Val Trp Gln Arg Lys Ser Lys Gly Lys	
20 25 30	
CTG CCT CCG GGA CCC ACC CCA TTG CCC TTC ATT GGA AAC TAC CTG CAG	144
Leu Pro Pro Gly Pro Thr Pro Phe Ile Gly Asn Tyr Leu Gln	
25 35 40 45	
CTG AAC ACA GAG CAG ATG TAC AAC TCC CTC ATG AAG ATC AGT GAG CGC	192
Leu Asn Thr Glu Gln Met Tyr Asn Ser Leu Met Lys Ile Ser Glu Arg	
50 55 60	
TAT GGC CCC GTG TTC ACC ATT CAC TTG GGG CCC CGG CGG GTC GTG GTG	240
Tyr Gly Pro Val Phe Thr Ile His Leu Gly Pro Arg Arg Val Val Val	
65 70 75 80	
CTG TGT GGA CAT GAT GCC GTC AGG GAG GCT CTG GTG GAC CAG GCT GAG	288
Leu Cys Gly His Asp Ala Val Arg Glu Ala Leu Val Asp Gln Ala Glu	
35 85 90 95	
GAG TTC AGC GGG CGA GGC GAG CAA GCC ACC TTC GAC TGG GTC TTC AAA	336
Glu Phe Ser Gly Arg Gly Glu Gln Ala Thr Phe Asp Trp Val Phe Lys	
100 105 110	
GGC TAT GGC GTG GTA TTC AGC AAC GGG GAG CGC GCC AAG CAG CTC CGG	384
Gly Tyr Gly Val Val Phe Ser Asn Gly Glu Arg Ala Lys Gln Leu Arg	
115 120 125	
CGC TTC TCC ATC GCC ACC CTG CGG GAC TTC GGG GTG GGC AAG CGA GGC	432
Arg Phe Ser Ile Ala Thr Leu Arg Asp Phe Gly Val Gly Lys Arg Gly	
130 135 140	
ATC GAG GAG CGC ATC CAG GAG GAG GCG GGC TTC CTC ATC GAC GCC CTC	480
Ile Glu Glu Arg Ile Gln Glu Glu Ala Gly Phe Leu Ile Asp Ala Leu	
145 150 155 160	

50

55

	CGG GGC ACT GGC GGC GCC AAT ATC GAT CCC ACC TTC TTC CTG AGC CGC	528
	Arg Gly Thr Gly Gly Ala Asn Ile Asp Pro Thr Phe Phe Leu Ser Arg	
	165 170 175	
5	ACA GTC TCC AAT GTC ATC AGC TCC ATT GTC TTT GGG GAC CGC TTT GAC	576
	Thr Val Ser Asn Val Ile Ser Ser Ile Val Phe Gly Asp Arg Phe Asp	
	180 185 190	
	TAT AAG GAC AAA GAG TTC CTG TCA CTG TTG CGC ATG ATG CTA GGA ATC	624
10	Tyr Lys Asp Lys Glu Phe Leu Ser Leu Leu Arg Met Met Leu Gly Ile	
	195 200 205	
	TTC CAG TTC ACG TCA ACC TCC ACG GGG CAG CTC TAT GAG ATG TTC TCT	672
	Phe Gln Phe Thr Ser Thr Ser Thr Gly Gln Leu Tyr Glu Met Phe Ser	
	210 215 220	
15	TCG GTG ATG AAA CAC CTG CCA GGA CCA CAG CAA CAG GCC TTT CAG TTG	720
	Ser Val Met Lys His Leu Pro Gly Pro Gln Gln Gln Ala Phe Gln Leu	
	225 230 235 240	
	CTG CAA GGG CTG GAG GAC TTC ATA GCC AAG AAG GTG GAG CAC AAC CAG	768
20	Leu Gln Gly Leu Glu Asp Phe Ile Ala Lys Lys Val Glu His Asn Gln	
	245 250 255	
	CGC ACG CTG GAT CCC AAT TCC CCA CGG GAC TTC ATT GAC TCC TTT CTC	816
	Arg Thr Leu Pro Asn Ser Pro Arg Asp Phe Ile Asp Ser Phe Leu	
	260 265 270	
25	ATC CGC ATG CAG GAG GAG GAG AAG AAC CCC AAC ACG GAG TTC TAC TTG	864
	Ile Arg Met Gln Glu Glu Glu Lys Asn Pro Asn Thr Glu Phe Tyr Leu	
	275 280 285	
	AAA AAC CTG GTG ATG ACC ACG TTG AAC CTC TTC ATT GGG GGC ACC GAG	912
30	Lys Asn Leu Val Met Thr Leu Asn Leu Phe Ile Gly Gly Thr Glu	
	290 295 300	
	ACC GTC AGC ACC ACC CTG CGC TAT GGC TTC TTG CTG CTC ATG AAG CAC	960
	Thr Val Ser Thr Thr Leu Arg Tyr Gly Phe Leu Leu Leu Met Lys His	
	305 310 315 320	
35	CCA GAG GTG GAG GCC AAG GTC CAT GAG GAG ATT GAC AGA GTG ATC GGC	1008
	Pro Glu Val Glu Ala Lys Val His Glu Glu Ile Asp Arg Val Ile Gly	
	325 330 335	
	AAG AAC CGG CAG CCC AAG TTT GAG GAC CGG GCC AAG ATG CCC TAC ATG	1056
40	Lys Asn Arg Gln Pro Lys Phe Glu Asp Arg Ala Lys Met Pro Tyr Met	
	340 345 350	
	GAG GCA GTG ATC CAC GAG ATC CAA AGA TTT GGA GAC GTG ATC CCC ATG	1104
	Glu Ala Val Ile His Glu Ile Gln Arg Phe Gly Asp Val Ile Pro Met	
	355 360 365	
45	AGT TTG GCC CGC AGA GTC AAA AAG GAC ACC AAG TTT CGG GAT TTC TTC	1152
	Ser Leu Ala Arg Arg Val Lys Lys Asp Thr Lys Phe Arg Asp Phe Phe	
	370 375 380	
	CTC CCT AAG GGC ACC GAA GTG TAC CCT ATG CTG GGC TCT GTG CTG AGA	1200
50	Leu Pro Lys Gly Thr Glu Val Tyr Pro Met Leu Gly Ser Val Leu Arg	
	385 390 395 400	
55		

	GAC CCC AGT TTC TTC TCC AAC CCC CAG GAC TTC AAT CCC CAG CAC TTC	1248
	Asp Pro Ser Phe Phe Ser Asn Pro Gln Asp Phe Asn Pro Gln His Phe	
	405 410 415	
5	CTG AAT GAG AAG GGG CAG TTT AAG AAG AGT GAT GCT TTT GTG CCC TTT	1296
	Leu Asn Glu Lys Gly Gln Phe Lys Lys Ser Asp Ala Phe Val Pro Phe	
	420 425 430	
	TCC ATC GGA AAG CGG AAC TGT TTC GGA GAA GGC CTG GCC AGA ATG GAG	1344
	Ser Ile Gly Lys Arg Asn Cys Phe Gly Glu Gly Leu Ala Arg Met Glu	
10	435 440 445	
	CTC TTT CTC TTC TTC ACC ACC GTC ATG CAG AAC TTC CGC CTC AAG TCC	1392
	Leu Phe Leu Phe Phe Thr Thr Val Met Gln Asn Phe Arg Leu Lys Ser	
	450 455 460	
15	TCC CAG TCA CCT AAG GAC ATT GAC GTG TCC CCC AAA CAC GTG GGC TTT	1440
	Ser Gln Ser Pro Lys Asp Ile Asp Val Ser Pro Lys His Val Gly Phe	
	465 470 475 480	
	GCC ACG ATC CCA CGA AAC TAC ACC ATG AGC TTC CTG CCC CGC	1482
	Ala Thr Ile Pro Arg Asn Tyr Thr Met Ser Phe Leu Pro Arg	
20	485 490	
	TGA	1485

(2) INFORMATION FOR SEQ ID NO: 18:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 494 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met	Leu	Ala	Ser	Gly	Met	Leu	Leu	Val	Ala	Leu	Leu	Val	Cys	Leu	Thr
1				5					10					15	
Val	Met	Val	Leu	Met	Ser	Val	Trp	Gln	Gln	Arg	Lys	Ser	Lys	Gly	Lys
			20					25					30		
Leu	Pro	Pro	Gly	Pro	Thr	Pro	Leu	Pro	Phe	Ile	Gly	Asn	Tyr	Leu	Gln
			35				40					45			
Leu	Asn	Thr	Glu	Gln	Met	Tyr	Asn	Ser	Leu	Met	Lys	Ile	Ser	Glu	Arg
	50				55						60				
Tyr	Gly	Pro	Val	Phe	Thr	Ile	His	Leu	Gly	Pro	Arg	Arg	Val	Val	Val
	65				70				75					80	
Leu	Cys	Gly	His	Asp	Ala	Val	Arg	Glu	Ala	Leu	Val	Asp	Gln	Ala	Glu
				85					90					95	
Glu	Phe	Ser	Gly	Arg	Gly	Glu	Gln	Ala	Thr	Phe	Asp	Trp	Val	Phe	Lys
			100				105						110		

50

55

Gly Tyr Gly Val Val Phe Ser Asn Gly Glu Arg Ala Lys Gln Leu Arg
 115 120 125
 5 Arg Phe Ser Ile Ala Thr Leu Arg Asp Phe Gly Val Gly Lys Arg Gly
 130 135 140
 Ile Glu Glu Arg Ile Gln Glu Glu Ala Gly Phe Leu Ile Asp Ala Leu
 145 150 155 160
 10 Arg Gly Thr Gly Gly Ala Asn Ile Asp Pro Thr Phe Phe Leu Ser Arg
 165 170 175
 Thr Val Ser Asn Val Ile Ser Ser Ile Val Phe Gly Asp Arg Phe Asp
 180 185 190
 15 Tyr Lys Asp Lys Glu Phe Leu Ser Leu Leu Arg Met Met Leu Gly Ile
 195 200 205
 Phe Gln Phe Thr Ser Thr Ser Thr Gly Gln Leu Tyr Glu Met Phe Ser
 210 215 220
 20 Ser Val Met Lys His Leu Pro Gly Pro Gln Gln Gln Ala Phe Gln Leu
 225 230 235 240
 Leu Gln Gly Leu Glu Asp Phe Ile Ala Lys Lys Val Glu His Asn Gln
 245 250 255
 25 Arg Thr Leu Asp Pro Asn Ser Pro Arg Asp Phe Ile Asp Ser Phe Leu
 260 265 270
 Ile Arg Met Gln Glu Glu Glu Lys Asn Pro Asn Thr Glu Phe Tyr Leu
 275 280 285
 30 Lys Asn Leu Val Met Thr Thr Leu Asn Leu Phe Ile Gly Gly Thr Glu
 290 295 300
 Thr Val Ser Thr Thr Leu Arg Tyr Gly Phe Leu Leu Leu Met Lys His
 305 310 315 320
 Pro Glu Val Glu Ala Lys Val His Glu Glu Ile Asp Arg Val Ile Gly
 325 330 335
 40 Lys Asn Arg Gln Pro Lys Phe Glu Asp Arg Ala Lys Met Pro Tyr Met
 340 345 350
 Glu Ala Val Ile His Glu Ile Gln Arg Phe Gly Asp Val Ile Pro Met
 355 360 365
 45 Ser Leu Ala Arg Arg Val Lys Lys Asp Thr Lys Phe Arg Asp Phe Phe
 370 375 380
 Leu Pro Lys Gly Thr Glu Val Tyr Pro Met Leu Gly Ser Val Leu Arg
 385 390 395 400
 50 Asp Pro Ser Phe Phe Ser Asn Pro Gln Asp Phe Asn Pro Gln His Phe
 405 410 415

55

Leu Asn Glu Lys Gly Gln Phe Lys Lys Ser Asp Ala Phe Val Pro Phe
 420 425 430
 Ser Ile Gly Lys Arg Asn Cys Phe Gly Glu Gly Leu Ala Arg Met Glu
 435 440 445
 Leu Phe Leu Phe Phe Thr Thr Val Met Gln Asn Phe Arg Leu Lys Ser
 450 455 460
 Ser Gln Ser Pro Lys Asp Ile Asp Val Ser Pro Lys His Val Gly Phe
 465 470 475 480
 Ala Thr Ile Pro Arg Asn Tyr Thr Met Ser Phe Leu Pro Arg
 485 490

(2) INFORMATION FOR SEQ ID NO: 19:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1476 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1473

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ATG GAA CTC AGC GTC CTC CTC TTC CTT GCA CTC CTC ACA GGA CTC TTG	48
Met Glu Leu Ser Val Leu Leu Phe Leu Ala Leu Leu Thr Gly Leu Leu	
1 5 10 15	
CTA CTC CTG GTT CAG CGC CAC CCT AAC ACC CAT GAC CGC CTC CCA CCA	96
Leu Leu Leu Val Gln Arg His Pro Asn Thr His Asp Arg Leu Pro Pro	
20 25 30	
GGG CCC CGC CCT CTG CCC CTT TTG GGA AAC CTT CTG CAG ATG GAT AGA	144
Gly Pro Arg Pro Leu Pro Leu Leu Gly Asn Leu Leu Gln Met Asp Arg	
35 40 45	
AGA GGC CTA CTC AAA TCC TTT CTG AGG TTC CGA GAG AAA TAT GGG GAC	192
Arg Gly Leu Leu Lys Ser Phe Leu Arg Phe Arg Glu Lys Tyr Gly Asp	
50 55 60	
GTC TTC ACG GTA CAC CTG GGA CCG AGG CCC GTG GTC ATG CTG TGT GGA	240
Val Phe Thr Val His Leu Gly Pro Arg Pro Val Val Met Leu Cys Gly	
65 70 75 80	
GTA GAG GCC ATA CGG GAG GCC CTT GTG GAC AAG GCT GAG GCC TTC TCT	288
Val Glu Ala Ile Arg Glu Ala Leu Val Asp Lys Ala Glu Ala Phe Ser	
85 90 95	
GGC CGG GGA AAA ATC GCC ATG GTC GAC CCA TTC TTC CGG GGA TAT GGT	336
Gly Arg Gly Lys Ile Ala Met Val Asp Pro Phe Phe Arg Gly Tyr Gly	
100 105 110	

50

55

	GTG ATC TTT GCC AAT GGA AAC CGC TGG AAG GTG CTT CGG CGA TTC TCT	384
	Val Ile Phe Ala Asn Gly Asn Arg Trp Lys Val Leu Arg Arg Phe Ser	
	115 120 125	
5	GTG ACC ACT ATG AGG GAC TTC GGG ATG GGA AAG CGG AGT GTG GAG GAG	432
	Val Thr Thr Met Arg Asp Phe Gly Met Gly Lys Arg Ser Val Glu Glu	
	130 135 140	
	CGG ATT CAG GAG GAG GCT CAG TGT CTG ATA GAG GAG CTT CGG AAA TCC	480
10	Arg Ile Gln Glu Glu Ala Gln Cys Leu Ile Glu Glu Leu Arg Lys Ser	
	145 150 155 160	
	AAG GGG GCC CTC ATG GAC CCC ACC TTC CTC TTC CAG TCC ATT ACC GCC	528
	Lys Gly Ala Leu Met Asp Pro Thr Phe Leu Phe Gln Ser Ile Thr Ala	
	165 170 175	
15	AAC ATC ATC TGC TCC ATC GTC TTT GGA AAA CGA TTC CAC TAC CAA GAT	576
	Asn Ile Ile Cys Ser Ile Val Phe Gly Lys Arg Phe His Tyr Gln Asp	
	180 185 190	
	CAA GAG TTC CTG AAG ATG CTG AAC TTG TTC TAC CAG ACT TTT TCA CTC	624
20	Gln Glu Phe Leu Lys Met Leu Asn Leu Phe Tyr Gln Thr Phe Ser Leu	
	195 200 205	
	ATC AGC TCT GTA TTC GGC CAG CTG TTT GAG CTC TTC TCT GGC TTC TTG	672
	Ile Ser Ser Val Phe Gly Gln Leu Phe Glu Leu Phe Ser Gly Phe Leu	
	210 215 220	
25	AAA TAC TTT CCT GGG GCA CAC AGG CAA GTT TAC AAA AAC CTG CAG GAA	720
	Lys Tyr Phe Pro Gly Ala His Arg Gln Val Tyr Lys Asn Leu Gln Glu	
	225 230 235 240	
	ATC AAT GCT TAC ATT GGC CAC AGT GTG GAG AAG CAC CGT GAA ACC CTG	768
30	Ile Asn Ala Tyr Ile Gly His Ser Val Glu Lys His Arg Glu Thr Leu	
	245 250 255	
	GAC CCC AGC GCC CCC AAG GAC CTC ATC GAC ACC TAC CTG CTC CAC ATG	816
	Asp Pro Ser Ala Pro Lys Asp Leu Ile Asp Thr Tyr Leu Leu His Met	
	260 265 270	
35	GAA AAA GAG AAA TCC AAC GCA CAC AGT GAA TTC AGC CAC CAG AAC CTC	864
	Glu Lys Glu Lys Ser Asn Ala His Ser Glu Phe Ser His Gln Asn Leu	
	275 280 285	
	AAC CTC AAC ACG CTC TCG CTC TTC TTT GCT GGC ACT GAG ACC ACC AGC	912
40	Asn Leu Asn Thr Leu Ser Leu Phe Phe Ala Gly Thr Glu Thr Thr Ser	
	290 295 300	
	ACC ACT CTC CGC TAC GGC TTC CTG CTC ATG CTC AAA TAC CCT CAT GTT	960
	Thr Thr Leu Arg Tyr Gly Phe Leu Leu Met Leu Lys Tyr Pro His Val	
	305 310 315 320	
45	GCA GAG AGA GTC TAC AGG GAG ATT GAA CAG GTG ATT GGC CCA CAT CGC	1008
	Ala Glu Arg Val Tyr Arg Glu Ile Glu Gln Val Ile Gly Pro His Arg	
	325 330 335	
	CCT CCA GAG CTT CAT GAC CGA GCC AAA ATG CCA TAC ACA GAG GCA GTC	1056
50	Pro Pro Glu Leu His Asp Arg Ala Lys Met Pro Tyr Thr Glu Ala Val	
	340 345 350	

ATC TAT GAG ATT CAG AGA TTT TCC GAC CTT CTC CCC ATG GGT GTG CCC 1104
 Ile Tyr Glu Ile Gln Arg Phe Ser Asp Leu Leu Pro Met Gly Val Pro
 355 360 365
 5 CAC ATT GTC ACC CAA CAC ACC AGC TTC CGA GGG TAC ATC ATC CCC AAG 1152
 His Ile Val Thr Gln His Thr Ser Phe Arg Gly Tyr Ile Ile Pro Lys
 370 375 380
 GAC ACA GAA GTA TTT CTC ATC CTG AGC ACT GCT CTC CAT GAC CCA CAC 1200
 Asp Thr Glu Val Phe Leu Ile Leu Ser Thr Ala Leu His Asp Pro His
 385 390 395 400
 10 TAC TTT GAA AAA CCA GAC GCC TTC AAT CCT GAC CAC TTT CTG GAT GCC 1248
 Tyr Phe Glu Lys Pro Asp Ala Phe Asn Pro Asp His Phe Leu Asp Ala
 405 410 415
 AAT GGG GCA CTG AAA AAG ACT GAA GCT TTT ATC CCC TTC TCC TTA GGG 1296
 Asn Gly Ala Leu Lys Lys Thr Glu Ala Phe Ile Pro Phe Ser Leu Gly
 420 425 430
 AAG CGG ATT TGT CTT GGT GAA GGC ATC GCC CGT GCG GAA TTG TTC CTC 1344
 Lys Arg Ile Cys Leu Gly Glu Gly Ile Ala Arg Ala Glu Leu Phe Leu
 435 440 445
 20 TTC TTC ACC ACC ATC CTC CAG AAC TTC TCC ATG GCC AGC CCC GTG GCC 1392
 Phe Phe Thr Thr Ile Leu Gln Asn Phe Ser Met Ala Ser Pro Val Ala
 450 455 460
 CCA GAA GAC ATC GAT CTG ACA CCC CAG GAG TGT GGT GTG GGC AAA ATA 1440
 Pro Glu Asp Ile Asp Leu Thr Pro Gln Glu Cys Gly Val Gly Lys Ile
 465 470 475 480
 CCC CCA ACA TAC CAG ATC CGC TTC CTG CCC CGC TGA 1476
 Pro Pro Thr Tyr Gln Ile Arg Phe Leu Pro Arg
 485 490
 30

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

40 Met Glu Leu Ser Val Leu Leu Phe Leu Ala Leu Leu Thr Gly Leu Leu
 1 5 10 15
 Leu Leu Leu Val Gln Arg His Pro Asn Thr His Asp Arg Leu Pro Pro
 20 25 30
 45 Gly Pro Arg Pro Leu Pro Leu Leu Gly Asn Leu Leu Gln Met Asp Arg
 35 40 45
 50
 55

Arg Gly Leu Leu Lys Ser Phe Leu Arg Phe Arg Glu Lys Tyr Gly Asp
 50 55 60
 Val Phe Thr Val His Leu Gly Pro Arg Pro Val Val Met Leu Cys Gly
 5 65 70 75 80
 Val Glu Ala Ile Arg Glu Ala Leu Val Asp Lys Ala Glu Ala Phe Ser
 85 90 95
 Gly Arg Gly Lys Ile Ala Met Val Asp Pro Phe Phe Arg Gly Tyr Gly
 10 100 105 110
 Val Ile Phe Ala Asn Gly Asn Arg Trp Lys Val Leu Arg Arg Phe Ser
 115 120 125
 Val Thr Thr Met Arg Asp Phe Gly Met Gly Lys Arg Ser Val Glu Glu
 15 130 135 140
 Arg Ile Gln Glu Glu Ala Gln Cys Leu Ile Glu Glu Leu Arg Lys Ser
 145 150 155 160
 Lys Gly Ala Leu Met Asp Pro Thr Phe Leu Phe Gln Ser Ile Thr Ala
 20 165 170 175
 Asn Ile Ile Cys Ser Ile Val Phe Gly Lys Arg Phe His Tyr Gln Asp
 180 185 190
 Gln Glu Phe Leu Lys Met Leu Asn Leu Phe Tyr Gln Thr Phe Ser Leu
 25 195 200 205
 Ile Ser Ser Val Phe Gly Gln Leu Phe Glu Leu Phe Ser Gly Phe Leu
 210 215 220
 Lys Tyr Phe Pro Gly Ala His Arg Gln Val Tyr Lys Asn Leu Gln Glu
 30 225 230 235 240
 Ile Asn Ala Tyr Ile Gly His Ser Val Glu Lys His Arg Glu Thr Leu
 245 250 255
 Asp Pro Ser Ala Pro Lys Asp Leu Ile Asp Thr Tyr Leu Leu His Met
 35 260 265 270
 Glu Lys Glu Lys Ser Asn Ala His Ser Glu Phe Ser His Gln Asn Leu
 275 280 285
 Asn Leu Asn Thr Leu Ser Leu Phe Phe Ala Gly Thr Glu Thr Thr Ser
 40 290 295 300
 Thr Thr Leu Arg Tyr Gly Phe Leu Leu Met Leu Lys Tyr Pro His Val
 305 310 315 320
 Ala Glu Arg Val Tyr Arg Glu Ile Glu Gln Val Ile Gly Pro His Arg
 45 325 330 335
 Pro Pro Glu Leu His Asp Arg Ala Lys Met Pro Tyr Thr Glu Ala Val
 340 345 350
 Ile Tyr Glu Ile Gln Arg Phe Ser Asp Leu Leu Pro Met Gly Val Pro
 50 355 360 365

His Ile Val Thr Gln His Thr Ser Phe Arg Gly Tyr Ile Ile Pro Lys
 370 375 380
 Asp Thr Glu Val Phe Leu Ile Leu Ser Thr Ala Leu His Asp Pro His
 385 390 395 400
 Tyr Phe Glu Lys Pro Asp Ala Phe Asn Pro Asp His Phe Leu Asp Ala
 405 410 415
 Asn Gly Ala Leu Lys Lys Thr Glu Ala Phe Ile Pro Phe Ser Leu Gly
 420 425 430
 Lys Arg Ile Cys Leu Gly Glu Gly Ile Ala Arg Ala Glu Leu Phe Leu
 435 440 445
 Phe Phe Thr Thr Ile Leu Gln Asn Phe Ser Met Ala Ser Pro Val Ala
 450 455 460
 Pro Glu Asp Ile Asp Leu Thr Pro Gln Glu Cys Gly Val Gly Lys Ile
 465 470 475 480
 Pro Pro Thr Tyr Gln Ile Arg Phe Leu Pro Arg
 485 490

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1473 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1470

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ATG GAA CCT TTT GTG GTC CTG GTG CTG TGT CTC TCT TTT ATG CTT CTC 48
 Met Glu Pro Phe Val Val Leu Val Leu Cys Leu Ser Phe Met Leu Leu
 1 5 10 15
 TTT TCA CTC TGG AGA CAG AGC TGT AGG AGA AGG AAG CTC CCT CCT GGC 96
 Phe Ser Leu Trp Arg Gln Ser Cys Arg Arg Arg Lys Leu Pro Pro Gly
 20 25 30
 CCC ACT CCT CTT CCT ATT ATT GGA AAT ATG CTA CAG ATA GAT GTT AAG 144
 Pro Thr Pro Leu Pro Ile Ile Gly Asn Met Leu Gln Ile Asp Val Lys
 35 40 45
 GAC ATC TGC AAA TCT TTC ACC AAT TTC TCA AAA GTC TAT GGT CCT GTG 192
 Asp Ile Cys Lys Ser Phe Thr Asn Phe Ser Lys Val Tyr Gly Pro Val
 50 55 60

EP 0 644 267 A2

	TTC ACC GTG TAT TTT GGC ATG AAT CCC ATA GTG GTG TTT CAT GGA TAT	240
	Phe Thr Val Tyr Phe Gly Met Asn Pro Ile Val Val Phe His Gly Tyr	
	65 70 75 80	
5	GAG GCA GTG AAG GAA GCC CTG ATT GAT AAT GGA GAG GAG TTT TCT GGA	288
	Glu Ala Val Lys Glu Ala Leu Ile Asp Asn Gly Glu Glu Phe Ser Gly	
	85 90 95	
	AGA GGC AAT TCC CCA ATA TCT CAA AGA ATT ACT AAA GGA CTT GGA ATC	336
10	Arg Gly Asn Ser Pro Ile Ser Gln Arg Ile Thr Lys Gly Leu Gly Ile	
	100 105 110	
	ATT TCC AGC AAT GGA AAG AGA TGG AAG GAG ATC CGG CGT TTC TCC CTC	384
	Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu	
	115 120 125	
15	ACA ACC TTG CGG AAT TTT GGG ATG GGG AAG AGG AGC ATT GAG GAC CGT	432
	Thr Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg	
	130 135 140	
	GTT CAA GAG GAA GCT CAC TGC CTT GTG GAG GAG TTG AGA AAA ACC AAG	480
20	Val Gln Glu Glu Ala His Cys Leu Val Glu Glu Leu Arg Lys Thr Lys	
	145 150 155 160	
	GCT TCA CCC TGT GAT CCC ACT TTC ATC CTG GGC TGT GCT CCC TGC AAT	528
	Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn	
	165 170 175	
25	GTG ATC TGC TCC GTT GTT TTC CAG AAA CGA TTT GAT TAT AAA GAT CAG	576
	Val Ile Cys Ser Val Val Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln	
	180 185 190	
	AAT TTT CTC ACC CTG ATG AAA AGA TTC AAT GAA AAC TTC AGG ATT CTG	624
30	Asn Phe Leu Thr Leu Met Lys Arg Phe Asn Glu Asn Phe Arg Ile Leu	
	195 200 205	
	AAC TCC CCA TGG ATC CAG GTC TGC AAT AAT TTC CCT CTA CTC ATT GAT	672
	Asn Ser Pro Trp Ile Gln Val Cys Asn Asn Phe Pro Leu Leu Ile Asp	
	210 215 220	
35	TGT TTC CCA GGA ACT CAC AAC AAA GTG CTT AAA AAT GTT GCT CTT ACA	720
	Cys Phe Pro Gly Thr His Asn Lys Val Leu Lys Asn Val Ala Leu Thr	
	225 230 235 240	
	CGA AGT TAC ATT AGG GAG AAA GTA AAA GAA CAC CAA GCA TCA CTG GAT	768
40	Arg Ser Tyr Ile Arg Glu Lys Val Lys Glu His Gln Ala Ser Leu Asp	
	245 250 255	
	GTT AAC AAT CCT CGG GAC TTT ATC GAT TGC TTC CTG ATC AAA ATG GAG	816
	Val Asn Asn Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu	
	260 265 270	
45	CAG GAA AAG GAC AAC CAA AAG TCA GAA TTC AAT ATT GAA AAC TTG GTT	864
	Gln Glu Lys Asp Asn Gln Lys Ser Glu Phe Asn Ile Glu Asn Leu Val	
	275 280 285	
	GGC ACT GTA GCT GAT CTA TTT GTT GCT GGA ACA GAG ACA ACA AGC ACC	912
50	Gly Thr Val Ala Asp Leu Phe Val Ala Gly Thr Glu Thr Thr Ser Thr	
	290 295 300	

55

	ACT CTG AGA TAT GGA CTC CTG CTC CTG CTG AAG CAC CCA GAG GTC ACA	960
	Thr Leu Arg Tyr Gly Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr	
	305 310 315 320	
5	GCT AAA GTC CAG GAA GAG ATT GAT CAT GTA ATT GGC AGA CAC AGG AGC	1008
	Ala Lys Val Gln Glu Glu Ile Asp His Val Ile Gly Arg His Arg Ser	
	325 330 335	
	CCC TGC ATG CAG GAT AGG AGC CAC ATG CCT TAC ACT GAT GCT GTA GTG	1056
	Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val	
	340 345 350	
10	CAC GAG ATC CAG AGA TAC AGT GAC CTT GTC CCC ACC GGT GTG CCC CAT	1104
	His Glu Ile Gln Arg Tyr Ser Asp Leu Val Pro Thr Gly Val Pro His	
	355 360 365	
15	GCA GTG ACC ACT GAT ACT AAG TTC AGA AAC TAC CTC ATC CCC AAG GGC	1152
	Ala Val Thr Thr Asp Thr Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly	
	370 375 380	
	ACA ACC ATA ATG GCA TTA CTG ACT TCC GTG CTA CAT GAT GAC AAA GAA	1200
	Thr Thr Ile Met Ala Leu Leu Thr Ser Val Leu His Asp Asp Lys Glu	
	385 390 395 400	
20	TTT CCT AAT CCA AAT ATC TTT GAC CCT GGC CAC TTT CTA GAT AAG AAT	1248
	Phe Pro Asn Pro Asn Ile Phe Asp Pro Gly His Phe Leu Asp Lys Asn	
	405 410 415	
25	GGC AAC TTT AAG AAA AGT GAC TAC TTC ATG CCT TTC TCA GCA GGA AAA	1296
	Gly Asn Phe Lys Lys Ser Asp Tyr Phe Met Pro Phe Ser Ala Gly Lys	
	420 425 430	
	CGA ATT TGT GCA GGA GAA GGA CTT GCC CGC ATG GAG CTA TTT TTA TTT	1344
	Arg Ile Cys Ala Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe	
	435 440 445	
30	CTA ACC ACA ATT TTA CAG AAC TTT AAC CTG AAA TCT GTT GAT GAT TTA	1392
	Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Val Asp Asp Leu	
	450 455 460	
35	AAG AAC CTC AAT ACT ACT GCA GTT ACC AAA GGG ATT GTT TCT CTG CCA	1440
	Lys Asn Leu Asn Thr Thr Ala Val Thr Lys Gly Ile Val Ser Leu Pro	
	465 470 475 480	
	CCC TCA TAC CAG ATC TGC TTC ATC CCT GTC TGA	1473
	Pro Ser Tyr Gln Ile Cys Phe Ile Pro Val	
	485 490	

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 490 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

5 Met Glu Pro Phe Val Val Leu Val Leu Cys Leu Ser Phe Met Leu Leu
 1 5 10 15
 Phe Ser Leu Trp Arg Gln Ser Cys Arg Arg Arg Lys Leu Pro Pro Gly
 20 25 30
 Pro Thr Pro Leu Pro Ile Ile Gly Asn Met Leu Gln Ile Asp Val Lys
 35 40 45
 10 Asp Ile Cys Lys Ser Phe Thr Asn Phe Ser Lys Val Tyr Gly Pro Val
 50 55 60
 Phe Thr Val Tyr Phe Gly Met Asn Pro Ile Val Val Phe His Gly Tyr
 65 70 75 80
 15 Glu Ala Val Lys Glu Ala Leu Ile Asp Asn Gly Glu Glu Phe Ser Gly
 85 90 95
 Arg Gly Asn Ser Pro Ile Ser Gln Arg Ile Thr Lys Gly Leu Gly Ile
 100 105 110
 20 Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu
 115 120 125
 Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg
 130 135 140
 25 Val Gln Glu Glu Ala His Cys Leu Val Glu Glu Leu Arg Lys Thr Lys
 145 150 155 160
 Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
 165 170 175
 30 Val Ile Cys Ser Val Val Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln
 180 185 190
 Asn Phe Leu Thr Leu Met Lys Arg Phe Asn Glu Asn Phe Arg Ile Leu
 195 200 205
 Asn Ser Pro Trp Ile Gln Val Cys Asn Asn Phe Pro Leu Leu Ile Asp
 210 215 220
 40 Cys Phe Pro Gly Thr His Asn Lys Val Leu Lys Asn Val Ala Leu Thr
 225 230 235 240
 Arg Ser Tyr Ile Arg Glu Lys Val Lys Glu His Gln Ala Ser Leu Asp
 245 250 255
 45 Val Asn Asn Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu
 260 265 270
 Gln Glu Lys Asp Asn Gln Lys Ser Glu Phe Asn Ile Glu Asn Leu Val
 275 280 285
 50 Gly Thr Val Ala Asp Leu Phe Val Ala Gly Thr Glu Thr Thr Ser Thr
 290 295 300

Thr Leu Arg Tyr Gly Leu Leu Leu Leu Leu Lys His Pro Glu Val Thr
 305 310 315 320
 Ala Lys Val Gln Glu Glu Ile Asp His Val Ile Gly Arg His Arg Ser
 5 325 330 335
 Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val
 340 345 350
 His Glu Ile Gln Arg Tyr Ser Asp Leu Val Pro Thr Gly Val Pro His
 10 355 360 365
 Ala Val Thr Thr Asp Thr Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly
 370 375 380
 Thr Thr Ile Met Ala Leu Leu Thr Ser Val Leu His Asp Asp Lys Glu
 15 385 390 395 400
 Phe Pro Asn Pro Asn Ile Phe Asp Pro Gly His Phe Leu Asp Lys Asn
 405 410 415
 Gly Asn Phe Lys Lys Ser Asp Tyr Phe Met Pro Phe Ser Ala Gly Lys
 20 420 425 430
 Arg Ile Cys Ala Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe
 435 440 445
 Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Val Asp Asp Leu
 25 450 455 460
 Lys Asn Leu Asn Thr Thr Ala Val Thr Lys Gly Ile Val Ser Leu Pro
 465 470 475 480
 Pro Ser Tyr Gln Ile Cys Phe Ile Pro Val
 30 485 490

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1473 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1470

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ATG GAA CCT TTT GTG GTC CTG GTG CTG TGT CTC TCT TTT ATG CTT CTC 48
 45 Met Glu Pro Phe Val Val Leu Val Leu Cys Leu Ser Phe Met Leu Leu
 1 5 10 15
 TTT TCA CTC TGG AGA CAG AGC TGT AGG AGA AGG AAG CTC CCT CCT GGC 96
 Phe Ser Leu Trp Arg Gln Ser Cys Arg Arg Arg Lys Leu Pro Pro Gly
 20 25 30

	CCC	ACT	CCT	CTT	CCT	ATT	ATT	GGA	AAT	ATG	CTA	CAG	ATA	GAT	GTT	AAG	144
	Pro	Thr	Pro	Leu	Pro	Ile	Ile	Gly	Asn	Met	Leu	Gln	Ile	Asp	Val	Lys	
			35					40					45				
5	GAC	ATC	TGC	AAA	TCT	TTC	ACC	AAT	TTC	TCA	AAA	GTC	TAT	GGT	CCT	GTG	192
	Asp	Ile	Cys	Lys	Ser	Phe	Thr	Asn	Phe	Ser	Lys	Val	Tyr	Gly	Pro	Val	
		50					55					60					
	TTC	ACC	GTG	TAT	TTT	GGC	ATG	AAT	CCC	ATA	GTG	GTG	TTT	CAT	GGA	TAT	240
	Phe	Thr	Val	Tyr	Phe	Gly	Met	Asn	Pro	Ile	Val	Val	Phe	His	Gly	Tyr	
10		65				70					75					80	
	GAG	GCA	GTG	AAG	GAA	GCC	CTG	ATT	GAT	AAT	GGA	GAG	GAG	TTT	TCT	GGA	288
	Glu	Ala	Val	Lys	Glu	Ala	Leu	Ile	Asp	Asn	Gly	Glu	Glu	Phe	Ser	Gly	
					85					90					95		
15	AGA	GGC	AAT	TCC	CCA	ATA	TCT	CAA	AGA	ATT	ACT	AAA	GGA	CTT	GGA	ATC	336
	Arg	Gly	Asn	Ser	Pro	Ile	Ser	Gln	Arg	Ile	Thr	Lys	Gly	Leu	Gly	Ile	
				100					105					110			
	ATT	TCC	AGC	AAT	GGA	AAG	AGA	TGG	AAG	GAG	ATC	CGG	CGT	TTC	TCC	CTC	384
	Ile	Ser	Ser	Asn	Gly	Lys	Arg	Trp	Lys	Glu	Ile	Arg	Arg	Phe	Ser	Leu	
			115					120					125				
20	ACA	ACC	TTG	CGG	AAT	TTT	GGG	ATG	GGG	AAG	AAG	AGC	ATT	GAG	GAC	CGT	432
	Thr	Thr	Leu	Arg	Asn	Phe	Gly	Met	Gly	Lys	Lys	Ser	Ile	Glu	Asp	Arg	
			130				135						140				
	GTT	CAA	GAG	GAA	GCT	CAC	TGC	CTT	GTG	GAG	GAG	TTG	AGA	AAA	ACC	AAG	480
25	Val	Gln	Glu	Glu	Ala	His	Cys	Leu	Val	Glu	Glu	Leu	Arg	Lys	Thr	Lys	
		145				150					155					160	
	GCT	TCA	CCC	TGT	GAT	CCC	ACT	TTC	ATC	CTG	GGC	TGT	GCT	CCC	TGC	AAT	528
	Ala	Ser	Pro	Cys	Asp	Pro	Thr	Phe	Ile	Leu	Gly	Cys	Ala	Pro	Cys	Asn	
					165					170					175		
30	GTG	ATC	TGC	TCC	GTT	GTT	TTC	CAG	AAA	CGA	TTT	GAT	TAT	AAA	GAT	CAG	576
	Val	Ile	Cys		Val	Val	Phe	Gln	Lys	Arg	Phe	Asp	Tyr	Lys	Asp	Gln	
				180					185					190			
	AAT	TTT	CTC	ACC	CTG	ATG	AAA	AGA	TTC	AAT	GAA	AAC	TTC	AGG	ATT	CTG	624
35	Asn	Phe	Leu	Thr	Leu	Met	Lys	Arg	Phe	Asn	Glu	Asn	Phe	Arg	Ile	Leu	
			195					200					205				
	AAC	TCC	CCA	TGG	ATC	CAG	GTC	TGC	AAT	AAT	TTC	CCT	CTA	CTC	ATT	GAT	672
	Asn	Ser	Pro	Trp	Ile	Gln	Val	Cys	Asn	Asn	Phe	Pro	Leu	Leu	Ile	Asp	
			210				215					220					
40	TGT	TTC	CCA	GGA	ACT	CAC	AAC	AAA	GTG	CTT	AAA	AAT	GTT	GCT	CTT	ACA	720
	Cys	Phe	Pro	Gly	Thr	His	Asn	Lys	Val	Leu	Lys	Asn	Val	Ala	Leu	Thr	
						230					235					240	
	CGA	AGT	TAC	ATT	AGG	GAG	AAA	GTA	AAA	GAA	CAC	CAA	GCA	TCA	CTG	GAT	768
45	Arg	Ser	Tyr	Ile	Arg	Glu	Lys	Val	Lys	Glu	His	Gln	Ala	Ser	Leu	Asp	
					245					250					255		

50

55

EP 0 644 267 A2

	GTT AAC AAT CCT CGG GAC TTT ATC GAT TGC TTC CTG ATC AAA ATG GAG	816
	Val Asn Asn Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu	
	260 265 270	
5	CAG GAA AAG GAC AAC CAA AAG TCA GAA TTC AAT ATT GAA AAC TTG GTT	864
	Gln Glu Lys Asp Asn Gln Lys Ser Glu Phe Asn Ile Glu Asn Leu Val	
	275 280 285	
	GGC ACT GTA GCT GAT CTA TTT GTT GCT GGA ACA GAG ACA ACA AGC ACC	912
	Gly Thr Val Ala Asp Leu Phe Val Ala Gly Thr Glu Thr Thr Ser Thr	
	290 295 300	
10	ACT CTG AGA TAT GGA CTC CTG CTC CTG CTG AAG CAC CCA GAG GTC ACA	960
	Thr Leu Arg Tyr Gly Leu Leu Leu Leu Lys His Pro Glu Val Thr	
	305 310 315 320	
15	GCT AAA GTC CAG GAA GAG ATT GAT CAT GTA ATT GGC AGA CAC AGG AGC	1008
	Ala Lys Val Gln Glu Ile Asp His Val Ile Gly Arg His Arg Ser	
	325 330 335	
	CCC TGC ATG CAG GAT AGG AGC CAC ATG CCT TAC ACT GAT GCT GTA GTG	1056
	Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val	
	340 345 350	
20	CAC GAG ATC CAG AGA TAC AGT GAC CTT GTC CCC ACC GGT GTG CCC CAT	1104
	His Glu Ile Gln Arg Tyr Ser Asp Leu Val Pro Thr Gly Val Pro His	
	355 360 365	
25	GCA GTG ACC ACT GAT ACT AAG TTC AGA AAC TAC CTC ATC CCC AAG GGC	1152
	Ala Val Thr Thr Asp Thr Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly	
	370 375 380	
	ACA ACC ATA ATG GCA TTA CTG ACT TCC GTG CTA CAT GAT GAC AGA GAA	1200
	Thr Thr Ile Met Ala Leu Leu Thr Ser Val Leu His Asp Asp Arg Glu	
	385 390 395 400	
30	TTT CCT AAT CCA AAT ATC TTT GAC CCT GGC CAC TTT CTA GAT AAG AAT	1248
	Phe Pro Asn Pro Asn Ile Phe Asp Pro Gly His Phe Leu Asp Lys Asn	
	405 410 415	
35	GGC AAC TTT AAG AAA AGT GAC TAC TTC ATG CCT TTC TCA GCA GGA AAA	1296
	Gly Asn Phe Lys Lys Ser Asp Tyr Phe Met Pro Phe Ser Ala Gly Lys	
	420 425 430	
	CGA ATT TGT GCA GGA GAA GGA CTT GCC CGC ATG GAG CTA TTT TTA TTT	1344
	Arg Ile Cys Ala Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe	
	435 440 445	
40	CTA ACC ACA ATT TTA CAG AAC TTT AAC CTG AAA TCT GTT GAT GAT TTA	1392
	Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Val Asp Asp Leu	
	450 455 460	
45	AAG AAC CTC AAT ACT ACT GCA GTT ACC AAA GGG ATT GTT TCT CTG CCA	1440
	Lys Asn Leu Asn Thr Thr Ala Val Thr Lys Gly Ile Val Ser Leu Pro	
	465 470 475 480	
	CCC TCA TAC CAG ATC TGC TTC ATC CCT GTC TGA	1473
	Pro Ser Tyr Gln Ile Cys Phe Ile Pro Val	
	485 490	
50		
55		

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 490 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Met Glu Pro Phe Val Val Leu Val Leu Cys Leu Ser Phe Met Leu Leu
 1 5 10 15
 Phe Ser Leu Trp Arg Gln Ser Cys Arg Arg Arg Lys Leu Pro Pro Gly
 15 20 25 30
 Pro Thr Pro Leu Pro Ile Ile Gly Asn Met Leu Gln Ile Asp Val Lys
 35 40 45
 Asp Ile Cys Lys Ser Phe Thr Asn Phe Ser Lys Val Tyr Gly Pro Val
 20 50 55 60
 Phe Thr Val Tyr Phe Gly Met Asn Pro Ile Val Val Phe His Gly Tyr
 65 70 75 80
 Glu Ala Val Lys Glu Ala Leu Ile Asp Asn Gly Glu Glu Phe Ser Gly
 25 85 90 95
 Arg Gly Asn Ser Pro Ile Ser Gln Arg Ile Thr Lys Gly Leu Gly Ile
 100 105 110
 Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu
 30 115 120 125
 Thr Thr Leu Arg Asn Phe Gly Met Gly Lys Lys Ser Ile Glu Asp Arg
 130 135 140
 Val Gln Glu Glu Ala His Cys Leu Val Glu Glu Leu Arg Lys Thr Lys
 35 145 150 155 160
 Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
 40 165 170 175
 Val Ile Cys Ser Val Val Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln
 180 185 190
 Asn Phe Leu Thr Leu Met Lys Arg Phe Asn Glu Asn Phe Arg Ile Leu
 45 195 200 205
 Asn Ser Pro Trp Ile Gln Val Cys Asn Asn Phe Pro Leu Leu Ile Asp
 210 215 220
 Cys Phe Pro Gly Thr His Asn Lys Val Leu Lys Asn Val Ala Leu Thr
 50 225 230 235 240

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1473 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1470

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

	ATG GAA CCT TTT GTG GTC CTG GTG CTG TGT CTC TCT TTT ATG CTT CTC	48
	Met Glu Pro Phe Val Val Leu Val Leu Cys Leu Ser Phe Met Leu Leu	
	1 5 10 15	
10	TTT TCA CTC TGG AGA CAG AGC TGT AGG AGA AGG AAG CTC CCT CCT GGC	96
	Phe Ser Leu Trp Arg Gln Ser Cys Arg Arg Arg Lys Leu Pro Pro Gly	
	20 25 30	
	CCC ACT CCT CTT CCT ATT ATT GGA AAT ATG CTA CAG ATA GAT GTT AAG	144
	Pro Thr Pro Leu Pro Ile Ile Gly Asn Met Leu Gln Ile Asp Val Lys	
15	35 40 45	
	GAC ATC TGC AAA TCT TTC ACC AAT TTC TCA AAA GTC TAT GGT CCT GTG	192
	Asp Ile Cys Lys Ser Phe Thr Asn Phe Ser Lys Val Tyr Gly Pro Val	
	50 55 60	
20	TTC ACC GTG TAT TTT GGC ATG AAT CCC ATA GTG GTG TTT CAT GGA TAT	240
	Phe Thr Val Tyr Phe Gly Met Asn Pro Ile Val Val Phe His Gly Tyr	
	65 70 75 80	
	GTG GCA GTG AAG GAA GCC CTG ATT GAT AAT GGA GAG GAG TTT TCT GGA	288
	Val Ala Val Lys Glu Ala Leu Ile Asp Asn Gly Glu Glu Phe Ser Gly	
25	85 90 95	
	AGA GGC AAT TCC CCA ATA TCT CAA AGA ATT ACT AAA GGA CTT GGA ATC	336
	Arg Gly Asn Ser Pro Ile Ser Gln Arg Ile Thr Lys Gly Leu Gly Ile	
	100 105 110	
30	ATT TCC AGC AAT GGA AAG AGA TGG AAG GAG ATC CGG CGT TTC TCC CTC	384
	Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu	
	115 120 125	
	ACA ACC TTG CGG AAT TTT GGG ATG GGG AAG AAG AGC ATT GAG GAC CGT	432
	Thr Thr Leu Arg Asn Phe Gly Met Gly Lys Lys Ser Ile Glu Asp Arg	
35	130 135 140	
	GTT CAA GAG GAA GCT CAC TGC CTT GTG GAG GAG TTG AGA AAA ACC AAG	480
	Val Gln Glu Glu Ala His Cys Leu Val Glu Glu Leu Arg Lys Thr Lys	
	145 150 155 160	
40	GCT TCA CCC TGT GAT CCC ACT TTC ATC CTG GGC TGT GCT CCC TGC AAT	528
	Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn	
	165 170 175	
	GTG ATC TGC TCC GTT GTT TTC CAG AAA CGA TTT GAT TAT AAA GAT CAG	576
	Val Ile Cys Ser Val Val Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln	
	180 185 190	
45	AAT TTT CTC ACC CTG ATG AAA AGA TTC AAT GAA AAC TTC AGG ATT CTG	624
	Asn Phe Leu Thr Leu Met Lys Arg Phe Asn Glu Asn Phe Arg Ile Leu	
	195 200 205	

50

55

	AAC	TCC	CCA	TGG	ATC	CAG	GTC	TGC	AAT	AAT	TTC	CCT	CTA	CTC	ATT	GAT	672
	Asn	Ser	Pro	Trp	Ile	Gln	Val	Cys	Asn	Asn	Phe	Pro	Leu	Leu	Ile	Asp	
	210						215					220					
5	TGT	TTC	CCA	GGA	ACT	CAC	AAC	AAA	GTG	CTT	AAA	AAT	GTT	GCT	CTT	ACA	720
	Cys	Phe	Pro	Gly	Thr	His	Asn	Lys	Val	Leu	Lys	Asn	Val	Ala	Leu	Thr	
	225					230					235					240	
	CGA	AGT	TAC	ATT	AGG	GAG	AAA	GTA	AAA	GAA	CAC	CAA	GCA	TCA	CTG	GAT	768
10	Arg	Ser	Tyr	Ile	Arg	Glu	Lys	Val	Lys	Glu	His	Gln	Ala	Ser	Leu	Asp	
					245					250					255		
	GTT	AAC	AAT	CCT	CGG	GAC	TTT	ATC	GAT	TGC	TTC	CTG	ATC	AAA	ATG	GAG	816
	Val	Asn	Asn	Pro	Arg	Asp	Phe	Ile	Asp	Cys	Phe	Leu	Ile	Lys	Met	Glu	
				260					265					270			
15	CAG	GAA	AAG	GAC	AAC	CAA	AAG	TCA	GAA	TTC	AAT	ATT	GAA	AAC	TTG	GTT	864
	Gln	Glu	Lys	Asp	Asn	Gln	Lys	Ser	Glu	Phe	Asn	Ile	Glu	Asn	Leu	Val	
			275					280					285				
	GGC	ACT	GTA	GCT	GAT	CTA	TTT	GTT	GCT	GGA	ACA	GAG	ACA	ACA	AGC	ACC	912
20	Gly	Thr	Val	Ala	Asp	Leu	Phe	Val	Ala	Gly	Thr	Glu	Thr	Thr	Ser	Thr	
	290						295					300					
	ACT	CTG	AGA	TAT	GGA	CTC	CTG	CTC	CTG	CTG	AAG	CAC	CCA	GAG	GTC	ACA	960
	Thr	Leu	Arg	Tyr	Gly	Leu	Leu	Leu	Leu	Leu	Lys	His	Pro	Glu	Val	Thr	
	305					310					315					320	
25	GCT	AAA	GTC	CAG	GAA	GAG	ATT	GAT	CAT	GTA	ATT	GGC	AGA	CAC	AGG	AGC	1008
	Ala	Lys	Val	Gln	Glu	Glu	Ile	Asp	His	Val	Ile	Gly	Arg	His	Arg	Ser	
					325					330					335		
	CCC	TGC	ATG	CAG	GAT	AGG	AGC	CAC	ATG	CCT	TAC	ACT	GAT	GCT	GTA	GTG	1056
30	Pro	Cys	Met	Gln	Asp	Arg	Ser	His	Met	Pro	Tyr	Thr	Asp	Ala	Val	Val	
				340					345				350				
	CAC	GAG	ATC	CAG	AGA	TAC	AGT	GAC	CTT	GTC	CCC	ACC	GGT	GTG	CCC	CAT	1104
	His	Glu	Ile	Gln	Arg	Tyr	Ser	Asp	Leu	Val	Pro	Thr	Gly	Val	Pro	His	
				355				360					365				
35	GCA	GTG	ACC	ACT	GAT	ACT	AAG	TTC	AGA	AAC	TAC	CTC	ATC	CCC	AAG	GGC	1152
	Ala	Val	Thr	Thr	Asp	Thr	Lys	Phe	Arg	Asn	Tyr	Leu	Ile	Pro	Lys	Gly	
		370					375					380					
	ACA	ACC	ATA	ATG	GCA	TTA	CTG	ACT	TCC	GTG	CTA	CAT	GAT	GAC	AGA	GAA	1200
40	Thr	Thr	Ile	Met	Ala	Leu	Leu	Thr	Ser	Val	Leu	His	Asp	Asp	Arg	Glu	
	385					390					395					400	
	TTT	CCT	AAT	CCA	AAT	ATC	TTT	GAC	CCT	GGC	CAC	TTT	CTA	GAT	AAG	AAT	1248
	Phe	Pro	Asn	Pro	Asn	Ile	Phe	Asp	Pro	Gly	His	Phe	Leu	Asp	Lys	Asn	
				405						410					415		
	GGC	AAC	TTT	AAG	AAA	AGT	GAC	TAC	TTC	ATG	CCT	TTC	TCA	GCA	GGA	AAA	1296
45	Gly	Asn	Phe	Lys	Lys	Ser	Asp	Tyr	Phe	Met	Pro	Phe	Ser	Ala	Gly	Lys	
				420					425				430				
	CGA	ATT	TGT	GCA	GGA	GAA	GGA	CTT	GCC	CGC	ATG	GAG	CTA	TTT	TTA	TTT	1344
50	Arg	Ile	Cys	Ala	Gly	Glu	Gly	Leu	Ala	Arg	Met	Glu	Leu	Phe	Leu	Phe	
			435					440					445				
55																	

CTA ACC ACA ATT TTA CAG AAC TTT AAC CTG AAA TCT GTT GAT GAT TTA 1392
 Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Val Asp Asp Leu
 450 455 460

5 AAG AAC CTC AAT ACT ACT GCA GTT ACC AAA GGG ATT GTT TCT CTG CCA 1440
 Lys Asn Leu Asn Thr Thr Ala Val Thr Lys Gly Ile Val Ser Leu Pro
 465 470 475 480

CCC TCA TAC CAG ATC TGC TTC ATC CCT GTC TGA 1473
 Pro Ser Tyr Gln Ile Cys Phe Ile Pro Val
 485 490

10

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 490 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

20 Met Glu Pro Phe Val Val Leu Val Leu Cys Leu Ser Phe Met Leu Leu
 1 5 10 15

Phe Ser Leu Trp Arg Gln Ser Cys Arg Arg Arg Lys Leu Pro Pro Gly
 20 25 30

25 Pro Thr Pro Leu Pro Ile Ile Gly Asn Met Leu Gln Ile Asp Val Lys
 35 40 45

Asp Ile Cys Lys Ser Phe Thr Asn Phe Ser Lys Val Tyr Gly Pro Val
 50 55 60

30 Phe Thr Val Tyr Phe Gly Met Asn Pro Ile Val Val Phe His Gly Tyr
 65 70 75 80

Val Ala Val Lys Glu Ala Leu Ile Asp Asn Gly Glu Glu Phe Ser Gly
 85 90 95

35 Arg Gly Asn Ser Pro Ile Ser Gln Arg Ile Thr Lys Gly Leu Gly Ile
 100 105 110

Ile Ser Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu
 115 120 125

40 Thr Thr Leu Arg Asn Phe Gly Met Gly Lys Lys Ser Ile Glu Asp Arg
 130 135 140

Val Gln Glu Glu Ala His Cys Leu Val Glu Glu Leu Arg Lys Thr Lys
 145 150 155 160

45 Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
 165 170 175

50

55

Val Ile Cys Ser Val Val Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln
 180 185 190
 5 Asn Phe Leu Thr Leu Met Lys Arg Phe Asn Glu Asn Phe Arg Ile Leu
 195 200 205
 Asn Ser Pro Trp Ile Gln Val Cys Asn Asn Phe Pro Leu Leu Ile Asp
 210 215 220
 10 Cys Phe Pro Gly Thr His Asn Lys Val Leu Lys Asn Val Ala Leu Thr
 225 230 235 240
 Arg Ser Tyr Ile Arg Glu Lys Val Lys Glu His Gln Ala Ser Leu Asp
 245 250 255
 15 Val Asn Asn Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu
 260 265 270
 Gln Glu Lys Asp Asn Gln Lys Ser Glu Phe Asn Ile Glu Asn Leu Val
 275 280 285
 20 Gly Thr Val Ala Asp Leu Phe Val Ala Gly Thr Glu Thr Thr Ser Thr
 290 295 300
 Thr Leu Arg Tyr Gly Leu Leu Leu Leu Lys His Pro Glu Val Thr
 305 310 315 320
 25 Ala Lys Val Gln Glu Glu Ile Asp His Val Ile Gly Arg His Arg Ser
 325 330 335
 Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val
 340 345 350
 30 His Glu Ile Gln Arg Tyr Ser Asp Leu Val Pro Thr Gly Val Pro His
 355 360 365
 Ala Val Thr Thr Asp Thr Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly
 370 375 380
 35 Thr Thr Ile Met Ala Leu Leu Thr Ser Val Leu His Asp Asp Arg Glu
 385 390 395 400
 Phe Pro Asn Pro Asn Ile Phe Asp Pro Gly His Phe Leu Asp Lys Asn
 405 410 415
 40 Gly Asn Phe Lys Lys Ser Asp Tyr Phe Met Pro Phe Ser Ala Gly Lys
 420 425 430
 Arg Ile Cys Ala Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe
 435 440 445
 45 Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Val Asp Asp Leu
 450 455 460
 Lys Asn Leu Asn Thr Thr Ala Val Thr Lys Gly Ile Val Ser Leu Pro
 465 470 475 480
 50 Pro Ser Tyr Gln Ile Cys Phe Ile Pro Val
 485 490

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1473 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1470

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

15	ATG GAT CCA GCT GTG GCT CTG GTG CTC TGT CTC TCC TGT TTG TTT CTC	48
	Met Asp Pro Ala Val Ala Leu Val Leu Cys Leu Ser Cys Leu Phe Leu	
	1 5 10 15	
	CTT TCA CTC TGG AGG CAG AGC TCT GGA AGA GGG AGG CTC CCG TCT GGC	96
	Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Arg Leu Pro Ser Gly	
	20 25 30	
20	CCC ACT CCT CTC CCG ATT ATT GGA AAT ATC CTG CAG TTA GAT GTT AAG	144
	Pro Thr Pro Leu Pro Ile Ile Gly Asn Ile Leu Gln Leu Asp Val Lys	
	35 40 45	
25	GAC ATG AGC AAA TCC TTA ACC AAT TTC TCA AAA GTC TAT GGC CCT GTG	192
	Asp Met Ser Lys Ser Leu Thr Asn Phe Ser Lys Val Tyr Gly Pro Val	
	50 55 60	
	TTC ACT GTG TAT TTT GGC CTG AAG CCC ATT GTG GTG TTG CAT GGA TAT	240
	Phe Thr Val Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr	
	65 70 75 80	
30	GAA GCA GTG AAG GAG GCC CTG ATT GAT CAT GGA GAG GAG TTT TCT GGA	288
	Glu Ala Val Lys Glu Ala Leu Ile Asp His Gly Glu Glu Phe Ser Gly	
	85 90 95	
35	AGA GGA AGT TTT CCA GTG GCT GAA AAA GTT AAC AAA GGA CTT GGA ATC	336
	Arg Gly Ser Phe Pro Val Ala Glu Lys Val Asn Lys Gly Leu Gly Ile	
	100 105 110	
	CTT TTC AGC AAT GGA AAG AGA TGG AAG GAG ATC CGG CGT TTC TGC CTC	384
	Leu Phe Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Cys Leu	
	115 120 125	
40	ATG ACT CTG CGG AAT TTT GGG ATG GGG AAG AGG AGC ATC GAG GAC CGT	432
	Met Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg	
	130 135 140	
45	GTT CAA GAG GAA GCC CGC TGC CTT GTG GAG GAG TTG AGA AAA ACC AAT	480
	Val Gln Glu Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Asn	
	145 150 155 160	

	GCC TCA CCC TGT GAT CCC ACT TTC ATC CTG GGC TGT GCT CCC TGC AAT	528
	Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn	
	165 170 175	
5	GTG ATC TGC TCT GTT ATT TTC CAT GAT CGA TTT GAT TAT AAA GAT CAG	576
	Val Ile Cys Ser Val Ile Phe His Asp Arg Phe Asp Tyr Lys Asp Gln	
	180 185 190	
	AGG TTT CTT AAC TTG ATG GAA AAA TTC AAT GAA AAC CTC AGG ATT CTG	624
10	Arg Phe Leu Asn Leu Met Glu Lys Phe Asn Glu Asn Leu Arg Ile Leu	
	195 200 205	
	AGC TCT CCA TGG ATC CAG GTC TGC AAT AAT TTC CCT GCT CTC ATC GAT	672
	Ser Ser Pro Trp Ile Gln Val Cys Asn Asn Phe Pro Ala Leu Ile Asp	
	210 215 220	
15	TAT CTC CCA GGA AGT CAT AAT AAA ATA GCT GAA AAT TTT GCT TAC ATT	720
	Tyr Leu Pro Gly Ser His Asn Lys Ile Ala Glu Asn Phe Ala Tyr Ile	
	225 230 235 240	
	AAA AGT TAT GTA TTG GAG AGA ATA AAA GAA CAT CAA GAA TCC CTG GAC	768
20	Lys Ser Tyr Val Leu Glu Arg Ile Lys Glu His Gln Glu Ser Leu Asp	
	245 250 255	
	ATG AAC AGT GCT CGG GAC TTT ATT GAT TGT TTC CTG ATC AAA ATG GAA	816
	Met Asn Ser Ala Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu	
	260 265 270	
25	CAG GAA AAG CAC AAT CAA CAG TCT GAA TTT ACT GTT GAA AGC TTG ATA	864
	Gln Glu Lys His Asn Gln Gln Ser Glu Phe Thr Val Glu Ser Leu Ile	
	275 280 285	
	GCC ACT GTA ACT GAT ATG TTT GGG GCT GGA ACA GAG ACA ACG AGC ACC	912
30	Ala Thr Val Thr Asp Met Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr	
	290 295 300	
	ACT CTG AGA TAT GGA CTC CTG CTC CTG CTG AAG TAC CCA GAG GTC ACA	960
	Thr Leu Arg Tyr Gly Leu Leu Leu Leu Lys Tyr Pro Glu Val Thr	
	305 310 315 320	
35	GCT AAA GTC CAG GAA GAG ATT GAA TGT GTA GTT GGC AGA AAC CGG AGC	1008
	Ala Lys Val Gln Glu Glu Ile Glu Cys Val Val Gly Arg Asn Arg Ser	
	325 330 335	
	CCC TGT ATG CAG GAC AGG AGT CAC ATG CCC TAC ACA GAT GCT GTG GTG	1056
40	Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val	
	340 345 350	
	CAC GAG ATC CAG AGA TAC ATT GAC CTC CTC CCC ACC AAC CTG CCC CAT	1104
	His Glu Ile Gln Arg Tyr Ile Asp Leu Leu Pro Thr Asn Leu Pro His	
	355 360 365	
45	GCA GTG ACC TGT GAT GTT AAA TTC AAA AAC TAC CTC ATC CCC AAG GGC	1152
	Ala Val Thr Cys Asp Val Lys Phe Lys Asn Tyr Leu Ile Pro Lys Gly	
	370 375 380	
	ACG ACC ATA ATA ACA TCC CTG ACT TCT GTG CTG CAC AAT GAC AAA GAA	1200
50	Thr Thr Ile Ile Thr Ser Leu Thr Ser Val Leu His Asn Asp Lys Glu	
	385 390 395 400	

TTC CCC AAC CCA GAG ATG TTT GAC CCT GGC CAC TTT CTG GAT AAG AGT 1248
 Phe Pro Asn Pro Glu Met Phe Asp Pro Gly His Phe Leu Asp Lys Ser
 405 410 415
 5 GGC AAC TTT AAG AAA AGT GAC TAC TTC ATG CCT TTC TCA GCA GGA AAA 1296
 Gly Asn Phe Lys Lys Ser Asp Tyr Phe Met Pro Phe Ser Ala Gly Lys
 420 425 430
 CGG ATG TGT ATG GGA GAG GGC CTG GCC CGC ATG GAG CTG TTT TTA TTC 1344
 Arg Met Cys Met Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe
 435 440 445
 10 CTG ACC ACC ATT TTG CAG AAC TTT AAC CTG AAA TCT CAG GTT GAC CCA 1392
 Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Gln Val Asp Pro
 450 455 460
 15 AAG GAT ATT GAC ATC ACC CCC ATT GCC AAT GCA TTT GGT CGT GTG CCA 1440
 Lys Asp Ile Asp Ile Thr Pro Ile Ala Asn Ala Phe Gly Arg Val Pro
 465 470 475 480
 CCC TTG TAC CAG CTC TGC TTC ATT CCT GTC TGA 1473
 Pro Leu Tyr Gln Leu Cys Phe Ile Pro Val
 485 490
 20

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 490 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

30 Met Asp Pro Ala Val Ala Leu Val Leu Cys Leu Ser Cys Leu Phe Leu
 1 5 10 15
 Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Arg Leu Pro Ser Gly
 20 25 30
 35 Pro Thr Pro Leu Pro Ile Ile Gly Asn Ile Leu Gln Leu Asp Val Lys
 35 40 45
 Asp Met Ser Lys Ser Leu Thr Asn Phe Ser Lys Val Tyr Gly Pro Val
 50 55 60
 40 Phe Thr Val Tyr Phe Gly Leu Lys Pro Ile Val Val Leu His Gly Tyr
 65 70 75 80
 Glu Ala Val Lys Glu Ala Leu Ile Asp His Gly Glu Glu Phe Ser Gly
 85 90 95
 45 Arg Gly Ser Phe Pro Val Ala Glu Lys Val Asn Lys Gly Leu Gly Ile
 100 105 110

50

55

Leu Phe Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Cys Leu
 115 120 125
 5 Met Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg
 130 135 140
 Val Gln Glu Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Asn
 145 150 155 160
 10 Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
 155 170 175
 Val Ile Cys Ser Val Ile Phe His Asp Arg Phe Asp Tyr Lys Asp Gln
 180 185 190
 15 Arg Phe Leu Asn Leu Met Glu Lys Phe Asn Glu Asn Leu Arg Ile Leu
 195 200 205
 Ser Ser Pro Trp Ile Gln Val Cys Asn Asn Phe Pro Ala Leu Ile Asp
 210 215 220
 20 Tyr Leu Pro Gly Ser His Asn Lys Ile Ala Glu Asn Phe Ala Tyr Ile
 225 230 235 240
 Lys Ser Tyr Val Leu Glu Arg Ile Lys Glu His Gln Glu Ser Leu Asp
 245 250 255
 25 Met Asn Ser Ala Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu
 260 265 270
 Gln Glu Lys His Asn Gln Gln Ser Glu Phe Thr Val Glu Ser Leu Ile
 275 280 285
 30 Ala Thr Val Thr Asp Met Phe Gly Ala Gly Thr Glu Thr Thr Ser Thr
 290 295 300
 Thr Leu Arg Tyr Gly Leu Leu Leu Leu Lys Tyr Pro Glu Val Thr
 305 310 315 320
 35 Ala Lys Val Gln Glu Glu Ile Glu Cys Val Val Gly Arg Asn Arg Ser
 325 330 335
 Pro Cys Met Gln Asp Arg Ser His Met Pro Tyr Thr Asp Ala Val Val
 340 345 350
 40 His Glu Ile Gln Arg Tyr Ile Asp Leu Leu Pro Thr Asn Leu Pro His
 355 360 365
 Ala Val Thr Cys Asp Val Lys Phe Lys Asn Tyr Leu Ile Pro Lys Gly
 370 375 380
 45 Thr Thr Ile Ile Thr Ser Leu Thr Ser Val Leu His Asn Asp Lys Glu
 385 390 395 400
 Phe Pro Asn Pro Glu Met Phe Asp Pro Gly His Phe Leu Asp Lys Ser
 405 410 415
 50 Gly Asn Phe Lys Lys Ser Asp Tyr Phe Met Pro Phe Ser Ala Gly Lys
 420 425 430

Arg Met Cys Met Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe
 435 440 445

5 Leu Thr Thr Ile Leu Gln Asn Phe Asn Leu Lys Ser Gln Val Asp Pro
 450 455 460

Lys Asp Ile Asp Ile Thr Pro Ile Ala Asn Ala Phe Gly Arg Val Pro
 465 470 475 480

10 Pro Leu Tyr Gln Leu Cys Phe Ile Pro Val
 485 490

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1473 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1470

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

25 ATG GAT CCT TTT GTG GTC CTT GTG CTC TGT CTC TCA TGT TTG CTT CTC 48
 Met Asp Pro Phe Val Val Leu Val Leu Cys Leu Ser Cys Leu Leu Leu
 1 5 10 15

CTT TCA CTC TGG AGA CAG AGC TCT GGG AGA GGA AAA CTC CCT CCT GGC 96
 Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly
 20 25 30

30 CCC ACT CCT CTC CCA GTG ATT GGA AAT ATC CTA CAG ATA GAT ATT AAG 144
 Pro Thr Pro Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Asp Ile Lys
 35 40 45

35 GAT GTC AGC AAA TCC TTA ACC AAT CTC TCA AAA ATC TAT GGC CCT GTG 192
 Asp Val Ser Lys Ser Leu Thr Asn Leu Ser Lys Ile Tyr Gly Pro Val
 50 55 60

TTC ACT CTG TAT TTT GGC CTC GAG CGC ATG GTG GTG CTG CAT GGA TAT 240
 Phe Thr Leu Tyr Phe Gly Leu Glu Arg Met Val Val Leu His Gly Tyr
 65 70 75 80

40 GAA GTG GTG AAG GAA GCC CTG ATT GAT CTT GGA GAG GAG TTT TCT GGA 288
 Glu Val Val Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly
 85 90 95

45 AGA GGC CAT TTC CCA CTG GCT GAA AGA GCT AAC AGA GGA TTT GGA ATC 336
 Arg Gly His Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile
 100 105 110

50

55

	GTT	TTC	AGC	AAT	GGA	AAG	AGA	TGG	AAG	GAG	ATC	CGG	CGT	TTC	TCC	CTC	384
	Val	Phe	Ser	Asn	Gly	Lys	Arg	Trp	Lys	Glu	Ile	Arg	Arg	Phe	Ser	Leu	
			115					120					125				
5	ATG	ACG	CTG	CGG	AAT	TTT	GGG	ATG	GGG	AAG	AGG	AGC	ATT	GAG	GAC	CGT	432
	Met	Thr	Leu	Arg	Asn	Phe	Gly	Met	Gly	Lys	Arg	Ser	Ile	Glu	Asp	Arg	
		130					135					140					
	GTT	CAA	GAG	GAA	GCC	CGC	TGC	CTT	GTG	GAG	GAG	TTG	AGA	AAA	ACC	AAG	480
	Val	Gln	Glu	Glu	Ala	Arg	Cys	Leu	Val	Glu	Glu	Leu	Arg	Lys	Thr	Lys	
		145				150					155					160	
10	GCT	TCA	CCC	TGT	GAT	CCC	ACT	TTC	ATC	CTG	GGC	TGT	GCT	CCC	TGC	AAT	528
	Ala	Ser	Pro	Cys	Asp	Pro	Thr	Phe	Ile	Leu	Gly	Cys	Ala	Pro	Cys	Asn	
				165						170					175		
	GTG	ATC	TGC	TCC	ATT	ATT	TTC	CAG	AAA	CGT	TTC	GAT	TAT	AAA	GAT	CAG	576
15	Val	Ile	Cys	Ser	Ile	Ile	Phe	Gln	Lys	Arg	Phe	Asp	Tyr	Lys	Asp	Gln	
				180					185					190			
	CAA	TTT	CTT	AAC	TTG	ATG	GAA	AAA	TTG	AAT	GAA	AAC	ATC	AGG	ATT	GTA	624
	Gln	Phe	Leu	Asn	Leu	Met	Glu	Lys	Leu	Asn	Glu	Asn	Ile	Arg	Ile	Val	
			195					200					205				
20	AGC	ACC	CCC	TGG	ATC	CAG	ATA	TGC	AAT	AAT	TTT	CCC	ACT	ATC	ATT	GAT	672
	Ser	Thr	Pro	Trp	Ile	Gln	Ile	Cys	Asn	Asn	Phe	Pro	Thr	Ile	Ile	Asp	
		210					215					220					
	TAT	TTC	CCG	GGA	ACC	CAT	AAC	AAA	TTA	CTT	AAA	AAC	CTT	GCT	TTT	ATG	720
25	Tyr	Phe	Pro	Gly	Thr	His	Asn	Lys	Leu	Leu	Lys	Asn	Leu	Ala	Phe	Met	
		225				230					235					240	
	GAA	AGT	GAT	ATT	TTG	GAG	AAA	GTA	AAA	GAA	CAC	CAA	GAA	TCG	ATG	GAC	768
	Glu	Ser	Asp	Ile	Leu	Glu	Lys	Val	Lys	Glu	His	Gln	Glu	Ser	Met	Asp	
				245					250						255		
30	ATC	AAC	AAC	CCT	CGG	GAC	TTT	ATT	GAT	TGC	TTC	CTG	ATC	AAA	ATG	GAG	816
	Ile	Asn	Asn	Pro	Arg	Asp	Phe	Ile	Asp	Cys	Phe	Leu	Ile	Lys	Met	Glu	
				260					265					270			
	AAG	GAA	AAG	CAA	AAC	CAA	CAG	TCT	GAA	TTC	ACT	ATT	GAA	AAC	TTG	GTA	864
35	Lys	Glu	Lys	Gln	Asn	Gln	Gln	Ser	Glu	Phe	Thr	Ile	Glu	Asn	Leu	Val	
			275					280					285				
	ATC	ACT	GCA	GCT	GAC	TTA	CTT	GGA	GCT	GGG	ACA	GAG	ACA	ACA	AGC	ACA	912
	Ile	Thr	Ala	Ala	Asp	Leu	Leu	Gly	Ala	Gly	Thr	Glu	Thr	Thr	Ser	Thr	
		290				295						300					
40	ACC	CTG	AGA	TAT	GCT	CTC	CTT	CTC	CTG	CTG	AAG	CAC	CCA	GAG	GTC	ACA	960
	Thr	Leu	Arg	Tyr	Ala	Leu	Leu	Leu	Leu	Leu	Lys	His	Pro	Glu	Val	Thr	
		305				310					315					320	
	GCT	AAA	GTC	CAG	GAA	GAG	ATT	GAA	CGT	GTC	GTT	GGC	AGA	AAC	CGG	AGC	1008
45	Ala	Lys	Val	Gln	Glu	Glu	Ile	Glu	Arg	Val	Val	Gly	Arg	Asn	Arg	Ser	
				325						330					335		
	CCC	TGC	ATG	CAG	GAC	AGG	GGC	CAC	ATG	CCC	TAC	ACA	GAT	GCT	GTG	GTG	1056
	Pro	Cys	Met	Gln	Asp	Arg	Gly	His	Met	Pro	Tyr	Thr	Asp	Ala	Val	Val	
				340					345					350			

50

55

EP 0 644 267 A2

	CAC GAG GTC CAG AGA TAC ATC GAC CTC ATC CCC ACC AGC CTG CCC CAT	1104
	His Glu Val Gln Arg Tyr Ile Asp Leu Ile Pro Thr Ser Leu Pro His	
	355 360 365	
5	GCA GTG ACC TGT GAC GTT AAA TTC AGA AAC TAC CTC ATT CCC AAG GGC	1152
	Ala Val Thr Cys Asp Val Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly	
	370 375 380	
	ACA ACC ATA TTA ACT TCC CTC ACT TCT GTG CTA CAT GAC AAC AAA GAA	1200
	Thr Thr Ile Leu Thr Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu	
	385 390 395 400	
10	TTC CCC AAC CCA GAG ATG TTT GAC CCT CGT CAC TTT CTG GAT GAA GGT	1248
	Phe Pro Asn Pro Glu Met Phe Asp Pro Arg His Phe Leu Asp Glu Gly	
	405 410 415	
15	GGA AAT TTT AAG AAA AGT AAC TAC TTC ATG CCT TTC TCA GCA GGA AAA	1296
	Gly Asn Phe Lys Lys Ser Asn Tyr Phe Met Pro Phe Ser Ala Gly Lys	
	420 425 430	
	CGG ATT TGT GTG GGA GAG GGC CTG GCC CGC ATG GAG CTG TTT TTA TTC	1344
	Arg Ile Cys Val Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe	
	435 440 445	
20	CTG ACC TTC ATT TTA CAG AAC TTT AAC CTG AAA TCT CTG ATT GAC CCA	1392
	Leu Thr Phe Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Ile Asp Pro	
	450 455 460	
25	AAG GAC CTT GAC ACA ACT CCT GTT GTC AAT GGA TTT GCT TCT GTC CCG	1440
	Lys Asp Leu Asp Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro	
	465 470 475 480	
	CCC TTC TAT CAG CTG TGC TTC ATT CCT GTC TGA	1473
	Pro Phe Tyr Gln Leu Cys Phe Ile Pro Val	
	485 490	

30

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 490 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

40	Met Asp Pro Phe Val Val Leu Val Leu Cys Leu Ser Cys Leu Leu Leu
	1 5 10 15
	Leu Ser Leu Trp Arg Gln Ser Ser Gly Arg Gly Lys Leu Pro Pro Gly
	20 25 30
45	Pro Thr Pro Leu Pro Val Ile Gly Asn Ile Leu Gln Ile Asp Ile Lys
	35 40 45

50

55

Asp Val Ser Lys Ser Leu Thr Asn Leu Ser Lys Ile Tyr Gly Pro Val
 50 55 60
 5 Phe Thr Leu Tyr Phe Gly Leu Glu Arg Met Val Val Leu His Gly Tyr
 65 70 75 80
 Glu Val Val Lys Glu Ala Leu Ile Asp Leu Gly Glu Glu Phe Ser Gly
 85 90 95
 10 Arg Gly His Phe Pro Leu Ala Glu Arg Ala Asn Arg Gly Phe Gly Ile
 100 105 110
 Val Phe Ser Asn Gly Lys Arg Trp Lys Glu Ile Arg Arg Phe Ser Leu
 115 120 125
 15 Met Thr Leu Arg Asn Phe Gly Met Gly Lys Arg Ser Ile Glu Asp Arg
 130 135 140
 Val Gln Glu Glu Ala Arg Cys Leu Val Glu Glu Leu Arg Lys Thr Lys
 145 150 155 160
 20 Ala Ser Pro Cys Asp Pro Thr Phe Ile Leu Gly Cys Ala Pro Cys Asn
 165 170 175
 Val Ile Cys Ser Ile Ile Phe Gln Lys Arg Phe Asp Tyr Lys Asp Gln
 180 185 190
 25 Gln Phe Leu Asn Leu Met Glu Lys Leu Asn Glu Asn Ile Arg Ile Val
 195 200 205
 Ser Thr Pro Trp Ile Gln Ile Cys Asn Asn Phe Pro Thr Ile Ile Asp
 210 215 220
 30 Tyr Phe Pro Gly Thr His Asn Lys Leu Leu Lys Asn Leu Ala Phe Met
 225 230 235 240
 Glu Ser Asp Ile Leu Glu Lys Val Lys Glu His Gln Glu Ser Met Asp
 245 250 255
 35 Ile Asn Asn Pro Arg Asp Phe Ile Asp Cys Phe Leu Ile Lys Met Glu
 260 265 270
 Lys Glu Lys Gln Asn Gln Gln Ser Glu Phe Thr Ile Glu Asn Leu Val
 275 280 285
 40 Ile Thr Ala Ala Asp Leu Leu Gly Ala Gly Thr Glu Thr Thr Ser Thr
 290 295 300
 Thr Leu Arg Tyr Ala Leu Leu Leu Leu Lys His Pro Glu Val Thr
 305 310 315 320
 45 Ala Lys Val Gln Glu Glu Ile Glu Arg Val Val Gly Arg Asn Arg Ser
 325 330 335
 Pro Cys Met Gln Asp Arg Gly His Met Pro Tyr Thr Asp Ala Val Val
 340 345 350
 50 His Glu Val Gln Arg Tyr Ile Asp Leu Ile Pro Thr Ser Leu Pro His
 355 360 365

Ala Val Thr Cys Asp Val Lys Phe Arg Asn Tyr Leu Ile Pro Lys Gly
 370 375 380

Thr Thr Ile Leu Thr Ser Leu Thr Ser Val Leu His Asp Asn Lys Glu
 385 390 395 400

Phe Pro Asn Pro Glu Met Phe Asp Pro Arg His Phe Leu Asp Glu Gly
 405 410 415

Gly Asn Phe Lys Lys Ser Asn Tyr Phe Met Pro Phe Ser Ala Gly Lys
 420 425 430

Arg Ile Cys Val Gly Glu Gly Leu Ala Arg Met Glu Leu Phe Leu Phe
 435 440 445

Leu Thr Phe Ile Leu Gln Asn Phe Asn Leu Lys Ser Leu Ile Asp Pro
 450 455 460

Lys Asp Leu Asp Thr Thr Pro Val Val Asn Gly Phe Ala Ser Val Pro
 465 470 475 480

Pro Phe Tyr Gln Leu Cys Phe Ile Pro Val
 485 490

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1494 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1491

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

ATG GGG CTA GAA GCA CTG GTG CCC CTG GCC GTG ATA GTG GCC ATC TTC 48
 Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe
 1 5 10 15

CTG CTC CTG GTG GAC CTG ATG CAC CGG CGC CAA CGC TGG GCT GCA CGC 96
 Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg
 20 25 30

TAC CCA CCA GGC CCC CTG CCA CTG CCC GGG CTG GGC AAC CTG CTG CAT 144
 Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His
 35 40 45

GTG GAC TTC CAG AAC ACA CCA TAC TGC TTC GAC CAG TTG CGG CGC CGC 192
 Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg
 50 55 60

	TTC	GGG	GAC	GTG	TTC	AGC	CTG	CAG	CTG	GCC	TGG	ACG	CCG	GTG	GTC	GTG	240
	Phe	Gly	Asp	Val	Phe	Ser	Leu	Gln	Leu	Ala	Trp	Thr	Pro	Val	Val	Val	
	65					70					75					80	
5	CTC	AAT	GGG	CTG	GCG	GCC	GTG	CGC	GAG	GCG	CTG	GTG	ACC	CAC	GGC	GAG	288
	Leu	Asn	Gly	Leu	Ala	Ala	Val	Arg	Glu	Ala	Leu	Val	Thr	His	Gly	Glu	
					85					90					95		
	GAC	ACC	GCC	GAC	CGC	CCG	CCT	GTG	CCC	ATC	ACC	CAG	ATC	CTG	GGT	TTC	336
10	Asp	Thr	Ala	Asp	Arg	Pro	Pro	Val	Pro	Ile	Thr	Gln	Ile	Leu	Gly	Phe	
				100					105					110			
	GGG	CCG	CGT	TCC	CAA	GGG	GTG	TTC	CTG	GCG	CGC	TAT	GGG	CCC	GCG	TGG	384
	Gly	Pro	Arg	Ser	Gln	Gly	Val	Phe	Leu	Ala	Arg	Tyr	Gly	Pro	Ala	Trp	
			115				120						125				
15	CGC	GAG	CAG	AGG	CGC	TTC	TCC	GTC	TCC	ACC	TTG	CGC	AAC	TTG	GGC	CTG	432
	Arg	Glu	Gln	Arg	Arg	Phe	Ser	Val	Ser	Thr	Leu	Arg	Asn	Leu	Gly	Leu	
		130					135					140					
	GGC	AAG	AAG	TCG	CTG	GAG	CAG	TGG	GTG	ACC	GAG	GAG	GCC	GCC	TGC	CTT	480
20	Gly	Lys	Lys	Ser	Leu	Glu	Gln	Trp	Val	Thr	Glu	Glu	Ala	Ala	Cys	Leu	
	145					150					155				160		
	TGT	GCC	GCC	TTC	GCC	AAC	CAC	TCC	GGA	CGC	CCC	TTT	CGC	CCC	AAC	GGT	528
	Cys	Ala	Ala	Phe	Ala	Asn	His	Ser	Gly	Arg	Pro	Phe	Arg	Pro	Asn	Gly	
					165					170					175		
25	CTC	TTG	GAC	AAA	GCC	GTG	AGC	AAC	GTG	ATC	GCC	TCC	CTC	ACC	TGC	GGG	576
	Leu	Leu	Asp	Lys	Ala	Val	Ser	Asn	Val	Ile	Ala	Ser	Leu	Thr	Cys	Gly	
				180					185					190			
	CGC	CGC	TTC	GAA	TAC	GAC	GAC	CCT	CGC	TTC	CTC	AGG	CTG	CTG	GAC	CTA	624
30	Arg	Arg	Phe	Glu	Tyr	Asp	Asp	Pro	Arg	Phe	Leu	Arg	Leu	Leu	Asp	Leu	
			195					200					205				
	GCT	CAG	GAG	GGA	CTG	AAG	GAG	GAG	TCG	GGC	TTT	CTG	CGC	GAG	GTG	CTG	672
	Ala	Gln	Glu	Gly	Leu	Lys	Glu	Glu	Ser	Gly	Phe	Leu	Arg	Glu	Val	Leu	
		210					215					220					
35	AAT	GCT	GTC	CCC	GTC	CTC	CTG	CAT	ATC	CCA	GCG	CTG	GCT	GGC	AAG	GTC	720
	Asn	Ala	Val	Pro	Val	Leu	Leu	His	Ile	Pro	Ala	Leu	Ala	Gly	Lys	Val	
	225					230					235				240		
	CTA	CGC	TTC	CAA	AAG	GCT	TTC	CTG	ACC	CAG	CTG	GAT	GAG	CTG	CTA	ACT	768
40	Leu	Arg	Phe	Gln	Lys	Ala	Phe	Leu	Thr	Gln	Leu	Asp	Glu	Leu	Leu	Thr	
					245					250					255		
	GAG	CAC	AGG	ATG	ACC	TGG	GAC	CCA	GCC	CAG	CCC	CCC	CGA	GAC	CTG	ACT	816
	Glu	His	Arg	Met	Thr	Trp	Asp	Pro	Ala	Gln	Pro	Pro	Arg	Asp	Leu	Thr	
				260					265					270			
45	GAG	GCC	TTC	CTG	GCA	GAG	ATG	GAG	AAG	GCC	AAG	GGG	AAC	CCT	GAG	AGC	864
	Glu	Ala	Phe	Leu	Ala	Glu	Met	Glu	Lys	Ala	Lys	Gly	Asn	Pro	Glu	Ser	
			275					280					285				
	AGC	TTC	AAT	GAT	GAG	AAC	CTG	TGC	ATA	GTG	GTG	GCT	GAC	CTG	TTC	TCT	912
50	Ser	Phe	Asn	Asp	Glu	Asn	Leu	Cys	Ile	Val	Val	Ala	Asp	Leu	Phe	Ser	
		290					295					300					
55																	

	GCC GGG ATG GTG ACC ACC TCG ACC ACG CTG GCC TGG GGC CTC CTG CTC	960
	Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu	
	305 310 315 320	
5	ATG ATC CTA CAT CCG GAT GTG CAG CGC CGT GTC CAA CAG GAG ATC GAC	1008
	Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp	
	325 330 335	
	GAC GTG ATA GGG CAG GTG CGG CGA CCA GAG ATG GGT GAC CAG GCT CAC	1056
	Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His	
10	340 345 350	
	ATG CCC TAC ACC ACT GCC GTG ATT CAT GAG GTG CAG CGC TTT GGG GAC	1104
	Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp	
	355 360 365	
15	ATC GTC CCC CTG GGT GTG ACC CAT ATG ACA TCC CGT GAC ATC GAA GTA	1152
	Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val	
	370 375 380	
	CAG GGC TTC CGC ATC CCT AAG GGA ACG ACA CTC ATC ACC AAC CTG TCA	1200
	Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser	
20	385 390 395 400	
	TCG GTG CTG AAG GAT GAG GCC GTC TGG GAG AAG CCC TTC CGC TTC CAC	1248
	Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His	
	405 410 415	
25	CCC GAA CAC TTC CTG GAT GCC CAG GGC CAC TTT GTG AAG CCG GAG GCC	1296
	Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala	
	420 425 430	
	TTC CTG CCT TTC TCA GCA GGC CGC CGT GCA TGC CTC GGG GAG CCC CTG	1344
	Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu	
30	435 440 445	
	GCC CGC ATG GAG CTC TTC CTC TTC TTC ACC TCC CTG CTG CAG CAC TTC	1392
	Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe	
	450 455 460	
35	AGC TTC TCG GTG CCC ACT GGA CAG CCC CGG CCC AGC CAC CAT GGT GTC	1440
	Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val	
	465 470 475 480	
	TTT GCT TTC CTG GTG ACC CCA TCC CCC TAT GAG CTT TGT GCT GTG CCC	1488
	Phe Ala Phe Leu Val Thr Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro	
40	485 490 495	
	CGC TAG	1494
	Arg .	

45 (2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 497 amino acids
 (B) TYPE: amino acid

50

55

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe
 1 5 10 15
 Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg
 20 25 30
 Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His
 35 40 45
 Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg
 50 55 60
 Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val
 65 70 75 80
 Leu Asn Gly Leu Ala Ala Val Arg Glu Ala Leu Val Thr His Gly Glu
 85 90 95
 Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe
 100 105 110
 Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp
 115 120 125
 Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu
 130 135 140
 Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu
 145 150 155 160
 Cys Ala Ala Phe Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly
 165 170 175
 Leu Leu Asp Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly
 180 185 190
 Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu
 195 200 205
 Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu
 210 215 220
 Asn Ala Val Pro Val Leu Leu His Ile Pro Ala Leu Ala Gly Lys Val
 225 230 235 240
 Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr
 245 250 255
 Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr
 260 265 270
 Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser
 275 280 285

55

Ser Phe Asn Asp Glu Asn Leu Cys Ile Val Val Ala Asp Leu Phe Ser
 290 295 300
 5 Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu
 305 310 315 320
 Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp
 325 330 335
 10 Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His
 340 345 350
 Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp
 355 360 365
 15 Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val
 370 375 380
 Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser
 385 390 395 400
 20 Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His
 405 410 415
 Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala
 420 425 430
 25 Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu
 435 440 445
 30 Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe
 450 455 460
 Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val
 465 470 475 480
 35 Phe Ala Phe Leu Val Thr Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro
 485 490 495

Arg

40

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1494 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1491

50

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

5	ATG GGG CTA GAA GCA CTG GTG CCC CTG GCC GTG ATA GTG GCC ATC TTC Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe	43
	1 5 10 15	
	CTG CTC CTG GTG GAC CTG ATG CAC CGG CGC CAA CGC TGG GCT GCA CGC Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg	96
	20 25 30	
10	TAC CCA CCA GGC CCC CTG CCA CTG CCC GGG CTG GGC AAC CTG CTG CAT Tyr Pro Gly Pro Leu Pro Leu Gly Leu Gly Asn Leu Leu His	144
	35 40 45	
	GTG GAC TTC CAG AAC ACA CCA TAC TGC TTC GAC CAG TTG CGG CGC CGC Val Asp Phe Gln Asn Thr Tyr Cys Phe Asp Gln Leu Arg Arg Arg	192
15	50 55 60	
	TTC GGG GAC GTG TTC AGC CTG CAG CTG GCC TGG ACG CCG GTG GTC GTG Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val	240
	65 70 75 80	
20	CTC AAT GGG CTG GCG GCC GTG CGC GAG GCG CTG GTG ACC CAC GGC GAG Leu Asn Gly Leu Ala Val Arg Glu Ala Leu Val Thr His Gly Glu	288
	85 90 95	
	GAC ACC GCC GAC CGC CCG CCT GTG CCC ATC ACC CAG ATC CTG GGT TTC Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe	336
25	100 105 110	
	GGG CCG CGT TCC CAA GGG GTG TTC CTG GCG CGC TAT GGG CCC GCG TGG Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp	384
	115 120 125	
30	CGC GAG CAG AGG CGC TTC TCC GTC TCC ACC TTG CGC AAC TTG GGC CTG Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu	432
	130 135 140	
	GGC AAG AAG TCG CTG GAG CAG TGG GTG ACC GAG GAG GCC GCC TGC CTT Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu	480
35	145 150 155 160	
	TGT GCC GCC TTC GCC AAC CAC TCC GGA CGC CCC TTT CGC CCC AAC GGT Cys Ala Ala Phe Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly	528
	165 170 175	
40	CTC TTG GAC AAA GCC GTG AGC AAC GTG ATC GCC TCC CTC ACC TGC GGG Leu Leu Asp Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly	576
	180 185 190	
	CGC CGC TTC GAA TAC GAC GAC CCT CGC TTC CTC AGG CTG CTG GAC CTA Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu	624
	195 200 205	
45	GCT CAG GAG GGA CTG AAG GAG GAG TCG GGC TTT CTG CGC GAG GTG CTG Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu	672
	210 215 220	

50

55

	AAT GCT GTC CCC GTC CTC CTG CAT ATC CCA GCG CTG GCT GGC AAG GTC	720
	Asn Ala Val Pro Val Leu Leu His Ile Pro Ala Leu Ala Gly Lys Val	
	225 230 235 240	
5	CTA CGC TTC CAA AAG GCT TTC CTG ACC CAG CTG GAT GAG CTG CTA ACT	768
	Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr	
	245 250 255	
	GAG CAC AGG ATG ACC TGG GAC CCA GCC CAG CCC CCC CGA GAC CTG ACT	816
	Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr	
	260 265 270	
10	GAG GCC TTC CTG GCA GAG ATG GAG AAG GCC AAG GGG AAC CCT GAG AGC	864
	Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser	
	275 280 285	
15	AGC TTC AAT GAT GAG AAC CTG CGC ATA GTG GTG GCT GAC CTG TTC TCT	912
	Ser Phe Asn Asp Glu Asn Leu Arg Ile Val Val Ala Asp Leu Phe Ser	
	290 295 300	
	GCC GGG ATG GTG ACC ACC TCG ACC ACG CTG GCC TGG GGC CTC CTG CTC	960
	Ala Gly Met Val Thr Thr Ser Thr Leu Ala Trp Gly Leu Leu Leu	
	305 310 315 320	
20	ATG ATC CTA CAT CCG GAT GTG CAG CGC CGT GTC CAA CAG GAG ATC GAC	1008
	Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp	
	325 330 335	
	GAC GTG ATA GGG CAG GTG CGG CGA CCA GAG ATG GGT GAC CAG GCT CAC	1056
25	Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His	
	340 345 350	
	ATG CCC TAC ACC ACT GCC GTG ATT CAT GAG GTG CAG CGC TTT GGG GAC	1104
	Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp	
	355 360 365	
30	ATC GTC CCC CTG GGT GTG ACC CAT ATG ACA TCC CGT GAC ATC GAA GTA	1152
	Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val	
	370 375 380	
	CAG GGC TTC CGC ATC CCT AAG GGA ACG ACA CTC ATC ACC AAC CTG TCA	1200
35	Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser	
	385 390 395 400	
	TCG GTG CTG AAG GAT GAG GCC GTC TGG GAG AAG CCC TTC CGC TTC CAC	1248
	Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His	
	405 410 415	
40	CCC GAA CAC TTC CTG GAT GCC CAG GGC CAC TTT GTG AAG CCG GAG GCC	1296
	Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala	
	420 425 430	
	TTC CTG CCT TTC TCA GCA GGC CGC CGT GCA TGC CTC GGG GAG CCC CTG	1344
45	Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu	
	435 440 445	
	GCC CGC ATG GAG CTC TTC CTC TTC TTC ACC TCC CTG CTG CAG CAC TTC	1392
	Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe	
	450 455 460	

50

55

AGC TTC TCG GTG CCC ACT GGA CAG CCC CGG CCC AGC CAC CAT GGT GTC 1440
 Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val
 465 470 475 480
 5 TTT GCT TTC CTG GTG ACC CCA TCC CCC TAT GAG CTT TGT GCT GTG CCC 1488
 Phe Ala Phe Leu Val Thr Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro
 485 490 495
 CGC TAG 1494
 Arg
 10

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 497 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe
 1 5 10 15
 25 Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg
 20 25 30
 Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His
 35 40 45
 30 Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg
 50 55 60
 Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val
 65 70 75 80
 35 Leu Asn Gly Leu Ala Ala Val Arg Glu Ala Leu Val Thr His Gly Glu
 85 90 95
 Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe
 100 105 110
 40 Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp
 115 120 125
 Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu
 130 135 140
 45 Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu
 145 150 155 160
 Cys Ala Ala Phe Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly
 165 170 175

50

55

Leu Leu Asp Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly
 180 185 190
 Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu
 195 200 205
 Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu
 210 215 220
 Asn Ala Val Pro Val Leu Leu His Ile Pro Ala Leu Ala Gly Lys Val
 225 230 235 240
 Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr
 245 250 255
 Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr
 260 265 270
 Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser
 275 280 285
 Ser Phe Asn Asp Glu Asn Leu Arg Ile Val Val Ala Asp Leu Phe Ser
 290 295 300
 Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu
 305 310 315 320
 Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp
 325 330 335
 Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His
 340 345 350
 Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp
 355 360 365
 Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val
 370 375 380
 Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser
 385 390 395 400
 Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His
 405 410 415
 Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala
 420 425 430
 Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu
 435 440 445
 Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe
 450 455 460
 Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val
 465 470 475 480

Phe Ala Phe Leu Val Thr Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro
 485 490 495

Arg

5

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1494 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1491

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

20	ATG GGG CTA GAA GCA CTG GTG CCC CTG GCC GTG ATA GTG GCC ATC TTC	48
	Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe	
	1 5 10 15	
	CTG CTC CTG GTG GAC CTG ATG CAC CGG CGC CAA CGC TGG GCT GCA CGC	96
	Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg	
	20 25 30	
25	TAC CCA CCA GGC CCC CTG CCA CTG CCC GGG CTG GGC AAC CTG CTG CAT	144
	Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His	
	35 40 45	
30	GTG GAC TTC CAG AAC ACA CCA TAC TGC TTC GAC CAG TTG CGG CGC CGC	192
	Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg	
	50 55 60	
	TTC GGG GAC GTG TTC AGC CTG CAG CTG GCC TGG ACG CCG GTG GTC GTG	240
	Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val	
	65 70 75 80	
35	CTC AAT GGG CTG GCG GCC GTG CGC GAG GCG CTG GTG ACC CAC GGC GAG	288
	Leu Asn Gly Leu Ala Ala Val Arg Glu Ala Leu Val Thr His Gly Glu	
	85 90 95	
40	GAC ACC GCC GAC CGC CCG CCT GTG CCC ATC ACC CAG ATC CTG GGT TTC	336
	Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe	
	100 105 110	
	GGG CCG CGT TCC CAA GGG GTG TTC CTG GCG CGC TAT GGG CCC GCG TGG	384
	Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp	
	115 120 125	
45	CGC GAG CAG AGG CGC TTC TCC GTC TCC ACC TTG CGC AAC TTG GGC CTG	432
	Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu	
	130 135 140	

50

55

	GGC AAG AAG TCG CTG GAG CAG TGG GTG ACC GAG GAG GCC GCC TGC CTT	480
	Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu	
	145 150 155 160	
5	TGT GCC GCC TTC GCC AAC CAC TCC GGA CGC CCC TTT CGC CCC AAC GGT	528
	Cys Ala Ala Phe Lys Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly	
	165 170 175	
	CTC TTG GAC AAA GCC GTG AGC AAC GTG ATC GCC TCC CTC ACC TGC GGG	576
	Leu Leu Asp Phe Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly	
10	180 185 190	
	CGC CGC TTC GAA TAC GAC GAC CCT CGC TTC CTC AGG CTG CTG GAC CTA	624
	Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu	
	195 200 205	
15	GCT CAG GAG GGA CTG AAG GAG GAG TCG GGC TTT CTG CGC GAG GTG CTG	672
	Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu	
	210 215 220	
	AAT GCT GTC CCC GTC CTC CTG CAT ATC CCA GCG CTG GCT GGC AAG GTC	720
	Asn Ala Val Pro Val Leu His Ile Pro Ala Leu Ala Gly Lys Val	
20	225 230 235 240	
	CTA CGC TTC CAA AAG GCT TTC CTG ACC CAG CTG GAT GAG CTG CTA ACT	768
	Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr	
	245 250 255	
25	GAG CAC AGG ATG ACC TGG GAC CCA GCC CAG CCC CCC CGA GAC CTG ACT	816
	Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr	
	260 265 270	
	GAG GCC TTC CTG GCA GAG ATG GAG AAG GCC AAG GGG AAC CCT GAG AGC	864
	Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser	
30	275 280 285	
	AGC TTC AAT GAT GAG AAC CTG CGC ATA GTG GTG GCT GAC CTG TTC TCT	912
	Ser Phe Asn Asp Glu Asn Leu Arg Ile Val Val Ala Asp Leu Phe Ser	
	290 295 300	
35	GCC GGG ATG GTG ACC ACC TCG ACC ACG CTG GCC TGG GGC CTC CTG CTC	960
	Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu	
	305 310 315 320	
	ATG ATC CTA CAT CCG GAT GTG CAG CGC CGT GTC CAA CAG GAG ATC GAC	1008
	Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp	
	325 330 335	
40	GAC GTG ATA GGG CAG GTG CGG CGA CCA GAG ATG GGT GAC CAG GCT CAC	1056
	Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His	
	340 345 350	
45	ATG CCC TAC ACC ACT GCC GTG ATT CAT GAG GTG CAG CGC TTT GGG GAC	1104
	Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp	
	355 360 365	
	ATC GTC CCC CTG GGT GTG ACC CAT ATG ACA TCC CGT GAC ATC GAA GTA	1152
	Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val	
	370 375 380	

50

55

	CAG GGC TTC CGC ATC CCT AAG GGA ACG ACA CTC ATC ACC AAC CTG TCA	1200
	Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser	
	385 390 395 400	
5	TCG GTG CTG AAG GAT GAG GCC GTC TGG GAG AAG CCC TTC CGC TTC CAC	1248
	Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His	
	405 410 415	
	CCC GAA CAC TTC CTG GAT GCC CAG GGC CAC TTT GTG AAG CCG GAG GCC	1296
	Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala	
	420 425 430	
10	TTC CTG CCT TTC TCA GCA GGC CGC CGT GCA TGC CTC GGG GAG CCC CTG	1344
	Phe Leu Pro Phe Ser Ala Gly Arg Ala Cys Leu Gly Glu Pro Leu	
	435 440 445	
	GCC CGC ATG GAG CTC TTC CTC TTC TTC ACC TCC CTG CTG CAG CAC TTC	1392
15	Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe	
	450 455 460	
	AGC TTC TCG GTG CCC ACT GGA CAG CCC CGG CCC AGC CAC CAT GGT GTC	1440
	Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val	
	465 470 475 480	
20	TTT GCT TTC CTG GTG AGC CCA TCC CCC TAT GAG CTT TGT GCT GTG CCC	1488
	Phe Ala Phe Leu Val Ser Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro	
	485 490 495	
25	CGC TAG	1494
	Arg	

(2) INFORMATION FOR SEQ ID NO: 36:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 497 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

	Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe	
	1 5 10 15	
40	Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg	
	20 25 30	
	Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His	
	35 40 45	
45	Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg	
	50 55 60	
	Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val	
	65 70 75 80	

50

55

Leu Asn Gly Leu Ala Ala Val Arg Glu Ala Leu Val Thr His Gly Glu
 85 90 95
 5 Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe
 100 105 110
 Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp
 115 120 125
 10 Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu
 130 135 140
 Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu
 145 150 155 160
 15 Cys Ala Ala Phe Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly
 165 170 175
 Leu Leu Asp Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly
 180 185 190
 20 Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu
 195 200 205
 Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu
 210 215 220
 25 Asn Ala Val Pro Val Leu Leu His Ile Pro Ala Leu Ala Gly Lys Val
 225 230 235 240
 Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr
 245 250 255
 30 Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr
 260 265 270
 Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser
 275 280 285
 Ser Phe Asn Asp Glu Asn Leu Arg Ile Val Val Ala Asp Leu Phe Ser
 290 295 300
 40 Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu
 305 310 315 320
 Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp
 325 330 335
 45 Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His
 340 345 350
 Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp
 355 360 365
 50 Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val
 370 375 380

55

Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser
 385 390 395 400
 5 Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His
 405 410 415
 Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala
 420 425 430
 10 Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu
 435 440 445
 Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe
 450 455 460
 15 Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val
 465 470 475 480
 Phe Ala Phe Leu Val Ser Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro
 485 490 495
 Arg

20

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1494 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1491

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

35 ATG GGG CTA GAA GCA CTG GTG CCC CTG GCC GTG ATA GTG GCC ATC TTC 48
 Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe
 1 5 10 15
 CTG CTC CTG GTG GAC CTG ATG CAC CGG CGC CAA CGC TGG GCT GCA CGC 96
 Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg
 20 25 30
 40 TAC CCA CCA GGC CCC CTG CCA CTG CCC GGG CTG GGC AAC CTG CTG CAT 144
 Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His
 35 40 45
 45 GTG GAC TTC CAG AAC ACA CCA TAC TGC TTC GAC CAG TTG CGG TGC CGC 192
 Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg
 50 55 60
 50 TTC GGG GAC GTG TTC AGC CTG CAG CTG GCC TGG ACG CCG GTG GTC GTG 240
 Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val
 65 70 75 80

55

	CTC	AAT	GGG	CTG	GCG	GCC	GTG	CGC	GAG	GCG	CTG	GTG	ACC	CAC	GGC	GAG	288
	Leu	Asn	Gly	Leu	Ala	Ala	Val	Arg	Glu	Ala	Leu	Val	Thr	His	Gly	Glu	
				85					90						95		
5	GAC	ACC	GCC	GAC	CGC	CCG	CCT	GTG	CCC	ATC	ACC	CAG	ATC	CTG	GGT	TTC	336
	Asp	Thr	Ala	Asp	Arg	Pro	Pro	Val	Pro	Ile	Thr	Gln	Ile	Leu	Gly	Phe	
				100					105					110			
	GGG	CCG	CGT	TCC	CAA	GGG	GTG	TTC	CTG	GCG	CGC	TAT	GGG	CCC	GCG	TGG	384
	Gly	Pro	Arg	Ser	Gln	Gly	Val	Phe	Leu	Ala	Arg	Tyr	Gly	Pro	Ala	Trp	
			115					120					125				
10	CGC	GAG	CAG	AGG	CGC	TTC	TCC	GTC	TCC	ACC	TTG	CGC	AAC	TTG	GGC	CTG	432
	Arg	Glu	Gln	Arg	Arg	Phe	Ser	Val	Ser	Thr	Leu	Arg	Asn	Leu	Gly	Leu	
		130					135					140					
15	GGC	AAG	AAG	TCG	CTG	GAG	CAG	TGG	GTG	ACC	GAG	GAG	GCC	GCC	TGC	CTT	480
	Gly	Lys	Lys	Ser	Leu	Glu	Gln	Trp	Val	Thr	Glu	Glu	Ala	Ala	Cys	Leu	
		145				150					155				160		
	TGT	GCC	GCC	TTC	GCC	AAC	CAC	TCC	GGA	CGC	CCC	TTT	CGC	CCC	AAC	GGT	528
	Cys	Ala	Ala	Phe	Ala	Asn	His	Ser	Gly	Arg	Pro	Phe	Arg	Pro	Asn	Gly	
				165					170						175		
20	CTC	TTG	GAC	AAA	GCC	GTG	AGC	AAC	GTG	ATC	GCC	TCC	CTC	ACC	TGC	GGG	576
	Leu	Leu	Asp	Lys	Ala	Val	Ser	Asn	Val	Ile	Ala	Ser	Leu	Thr	Cys	Gly	
				180					185					190			
25	CGC	CGC	TTC	GAA	TAC	GAC	GAC	CCT	CGC	TTC	CTC	AGG	CTG	CTG	GAC	CTA	624
	Arg	Arg	Phe	Glu	Tyr	Asp	Asp	Pro	Arg	Phe	Leu	Arg	Leu	Leu	Asp	Leu	
			195					200					205				
	GCT	CAG	GAG	GGA	CTG	AAG	GAG	GAG	TCG	GGC	TTT	CTG	CGC	GAG	GTG	CTG	672
	Ala	Gln	Glu	Gly	Leu	Lys	Glu	Glu	Ser	Gly	Phe	Leu	Arg	Glu	Val	Leu	
		210					215					220					
30	AAT	GCT	GTC	CCC	GTC	CTC	CTG	CAT	ATC	CCA	GCG	CTG	GCT	GGC	AAG	GTC	720
	Asn	Ala	Val	Pro	Val	Leu	Leu	His	Ile	Pro	Ala	Leu	Ala	Gly	Lys	Val	
		225				230					235				240		
35	CTA	CGC	TTC	CAA	AAG	GCT	TTC	CTG	ACC	CAG	CTG	GAT	GAG	CTG	CTA	ACT	768
	Leu	Arg	Phe	Gln	Lys	Ala	Phe	Leu	Thr	Gln	Leu	Asp	Glu	Leu	Leu	Thr	
				245						250					255		
	GAG	CAC	AGG	ATG	ACC	TGG	GAC	CCA	GCC	CAG	CCC	CCC	CGA	GAC	CTG	ACT	816
	Glu	His	Arg	Met	Thr	Trp	Asp	Pro	Ala	Gln	Pro	Pro	Arg	Asp	Leu	Thr	
				260					265					270			
40	GAG	GCC	TTC	CTG	GCA	GAG	ATG	GAG	AAG	GCC	AAG	GGG	AAC	CCT	GAG	AGC	864
	Glu	Ala	Phe	Leu	Ala	Glu	Met	Glu	Lys	Ala	Lys	Gly	Asn	Pro	Glu	Ser	
			275					280					285				
45	AGC	TTC	AAT	GAT	GAG	AAC	CTG	TGC	ATA	GTG	GTG	GCT	GAC	CTG	TTC	TCT	912
	Ser	Phe	Asn	Asp	Glu	Asn	Leu	Cys	Ile	Val	Val	Ala	Asp	Leu	Phe	Ser	
		290					295					300					

50

55

	GCC GGG ATG GTG ACC ACC TCG ACC ACG CTG GCC TGG GGC CTC CTG CTC	960
	Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu	
	305 310 315 320	
5	ATG ATC CTA CAT CCG GAT GTG CAG CGC CGT GTC CAA CAG GAG ATC GAC	1008
	Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp	
	325 330 335	
	GAC GTG ATA GGG CAG GTG CGG CGA CCA GAG ATG GGT GAC CAG GCT CAC	1056
	Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His	
	340 345 350	
10	ATG CCC TAC ACC ACT GCC GTG ATT CAT GAG GTG CAG CGC TTT GGG GAC	1104
	Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp	
	355 360 365	
15	ATC GTC CCC CTG GGT GTG ACC CAT ATG ACA TCC CGT GAC ATC GAA GTA	1152
	Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val	
	370 375 380	
	CAG GGC TTC CGC ATC CCT AAG GGA ACG ACA CTC ATC ACC AAC CTG TCA	1200
	Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser	
	385 390 395 400	
20	TCG GTG CTG AAG GAT GAG GCC GTC TGG GAG AAG CCC TTC CGC TTC CAC	1248
	Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His	
	405 410 415	
25	CCC GAA CAC TTC CTG GAT GCC CAG GGC CAC TTT GTG AAG CCG GAG GCC	1296
	Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala	
	420 425 430	
	TTC CTG CCT TTC TCA GCA GGC CGC CGT GCA TGC CTC GGG GAG CCC CTG	1344
	Phe Leu Pro Phe Ser Ala Gly Arg Ala Cys Leu Gly Glu Pro Leu	
	435 440 445	
30	GCC CGC ATG GAG CTC TTC CTC TTC TTC ACC TCC CTG CTG CAG CAC TTC	1392
	Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe	
	450 455 460	
35	AGC TTC TCG GTG CCC ACT GGA CAG CCC CGG CCC AGC CAC CAT GGT GTC	1440
	Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val	
	465 470 475 480	
	TTT GCT TTC CTG GTG AGC CCA TCC CCC TAT GAG CTT TGT GCT GTG CCC	1488
	Phe Ala Phe Leu Val Ser Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro	
	485 490 495	
40	CGC TAG	1494
	Arg	

45 (2) INFORMATION FOR SEQ ID NO: 38:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 497 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50

55

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

5 Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe
 1 5 10 15
 Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg
 20 25 30
 10 Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His
 35 40 45
 Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg
 50 55 60
 15 Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val
 65 70 75 80
 Leu Asn Gly Leu Ala Ala Val Arg Glu Ala Leu Val Thr His Gly Glu
 85 90 95
 20 Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe
 100 105 110
 Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp
 115 120 125
 25 Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu
 130 135 140
 Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu
 145 150 155 160
 Cys Ala Ala Phe Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly
 165 170 175
 30 Leu Leu Asp Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly
 180 185 190
 Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu
 195 200 205
 40 Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu
 210 215 220
 Asn Ala Val Pro Val Leu Leu His Ile Pro Ala Leu Ala Gly Lys Val
 225 230 235 240
 45 Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr
 245 250 255
 Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr
 260 265 270
 50 Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser
 275 280 285

55

Ser Phe Asn Asp Glu Asn Leu Cys Ile Val Val Ala Asp Leu Phe Ser
 290 295 300
 Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu
 305 310 315 320
 Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp
 325 330 335
 Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His
 340 345 350
 Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp
 355 360 365
 Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val
 370 375 380
 Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser
 385 390 395 400
 Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His
 405 410 415
 Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala
 420 425 430
 Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu
 435 440 445
 Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe
 450 455 460
 Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val
 465 470 475 480
 Phe Ala Phe Leu Val Ser Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro
 485 490 495
 Arg

35
 (2) INFORMATION FOR SEQ ID NO: 39:

40
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

45
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

GGAACGCATG GTGGTGCTGC ATGGATATGA AGTG

34

50

55

(2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 56 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

CTCAAAGATC TATGGCCCTG TGTCTACTCT GTATTTTGGC CTCGAGCGCA TGGTGG

56

(2) INFORMATION FOR SEQ ID NO: 41:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

CCACCATGCG CTCGAGGCCA AAATACAG

28

(2) INFORMATION FOR SEQ ID NO: 42:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

GGGTTCCCGG GAAATAATCA ATGATAGTGG G

31

(2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

GGATTGTAAG CACCCCCTGG ATCCAGATAT GC

32

(2) INFORMATION FOR SEQ ID NO: 44:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

.CCCAGCTCCA AGTAAGTCAG CTGCAGTGAT TACC

34

(2) INFORMATION FOR SEQ ID NO: 45:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

GGTGGTACCC TTGGGAATGA GGTAGTTTCT GAATTTAACG TC

42

(2) INFORMATION FOR SEQ ID NO: 46:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

AGTCTAGAAT GGATCCTTTT GTGGTCCTTG TGC

33

(2) INFORMATION FOR SEQ ID NO: 47:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

CCCAGAGCTC TGTCTCCAGA GTGAAAGGAG

30

(2) INFORMATION FOR SEQ ID NO: 48:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

ACAGAGCTCT GGGAGAGGAA AACTCCCTCC

30

(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 54 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

CCATAGATTT TTGAGAGATT GGTTAAGGAT TTGCTGACAT CCTTAATATC TATC

54

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

25

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

GACCCTCGTC ACTTTCTGGA TGAAGGTGGA

35

30

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

40

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

GAAGTAGTTA CTTTCTTAA AATTTCCACC TTCATC

36

50

55

(2) INFORMATION FOR SEQ ID NO: 52:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

AAAGAATTCC CCAACCCAGA GATGTTTGAC CCTCGTC

37

(2) INFORMATION FOR SEQ ID NO: 53:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

GGCCAGGCC TCTCCACAC AAATCCGTTT TCCTGCTGAG AAAGGCATGA AGTAGTTAC

59

(2) INFORMATION FOR SEQ ID NO: 54:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

GAGAGGGCCT GGCCCGCATG GAGCTGTTTT TATTCCTGAC CTC

44

(2) INFORMATION FOR SEQ ID NO: 55:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

CAGGAGTTGT GTCAAGGTCC TTTGGGTCAA TCAG

34

(2) INFORMATION FOR SEQ ID NO: 56:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 64 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

TTGTCAATGG ATTTGCTTCT GTCCCGCCCT TCTATCAGCT GTGCTTCATT CCTGTCTGAG 60
 GATC 64

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 55 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

CAGAAGCAAA TCCATTGACA ACAGGAGTTG TGTCAAGGTC CTTTGGGTCA ATCAG 55

(2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 60 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

CTCAGACAGG AATGAAGCAC AGCTGATAGA AGGGCGGGAC AGAAGCAAAT CCATTGACAA 60

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

GCAGCCAGAC CATCTGTGCT TCTTCAGACA GG 32

(2) INFORMATION FOR SEQ ID NO: 60:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

CACCATATTA ACTTCCCTCA CTTCTGTGCT ACATGACAAC AAAG

44

(2) INFORMATION FOR SEQ ID NO: 61:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 52 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

AATTCTTTGT TGTCATGTAG CACAGAAGTG AGGGAAGTTA ATATGGTGGT AC

52

Claims

1. A method for evaluation of the safety of a chemical compound, which comprises the steps of:
 - (a) reacting a chemical compound with recombinant yeast cells producing human cytochrome P450 molecular species P450 1A2, P450 2C9, P450 2E1 and P450 3A4 together with a yeast NADPH-P450 reductase, which may be in the form of a fused enzyme with each of said human cytochrome P450 molecular species, or with the cell free extracts of the yeast cells; and
 - (b) analyzing the resulting metabolite to determine the safety of the compound.
2. The method according to claim 1, wherein the recombinant yeast cells are yeast cells transformed with plasmids having a gene coding for human cytochrome P450 1A2, P450 2C9, P450 2E1 or P450 3A4 together with a gene coding for yeast NADPH-P450 reductase.
3. The method according to claim 1 or 2, wherein the recombinant yeast cells are yeast cells transformed with plasmids each of which has a fused gene comprising a gene coding for the human cytochrome P450 molecule on the 5'-terminal and a gene coding for the yeast NADPH-P450 reductase on 3'-terminal.
4. The method according to any one of claims 1 to 3, wherein the analyzing of the metabolite is carried out by the Ames Test.
5. The method according to claim 4, wherein the test is carried out using His⁻ Salmonella or Trp⁻ Escherichia coli.
6. The method according to any one of claims 1 to 5, wherein the recombinant yeast cells further produce at least one additional human cytochrome P450 molecular species selected from a group of human cytochrome P450 2A6, P450 2C19 and P450 2D6.
7. The method according to any one of claims 1 to 6, wherein the recombinant yeast cells further produce at least one additional human cytochrome P450 molecular species selected from a group of human cytochrome P450 1A1, P450 2B6, P450 2C8 and P450 2C18.

8. An artificial fused enzyme, which comprises human cytochrome p450 3A4 and yeast NADPH-P450 reductase.
9. A yeast expression plasmid having a fused gene comprising a gene coding for human P450 3A4 and a
5 gene coding for the yeast NADPH-P450 reductase.
10. A method of determining in vitro the human metabolite of a chemical compound, which comprises the steps of:
10 (a) reacting a chemical compound with recombinant yeast cells producing human cytochrome P450 molecular species P450 1A2, P450 2C9, P450 2E1 and P450 3A4 together with a yeast NADPH-P450 reductase, which may be in the form of a fused enzyme with each of said human cytochrome P450 molecular species, or with the cell free extracts of the yeast cells; and
(b) identifying the resulting metabolite.

15

20

25

30

35

40

45

50

55

1A2	5' -CACAGAGCTCCTCCTGGCCCTCTGCCATCTTC-3'	Primer for amplifying P4501A2 1.5Kb fragment
	5' -TTACAGGCCCTGCACCTTGGCTAAGCTGC-3'	
2C9	5' -AGTCTAGAAATGGATTCATTTGTGTCCCTTGTGCTC-3'	Primer for amplifying P4502C9 0.9Kb fragment
	5' -CTCCAAACAAAGTCAACTGCAGTGTTTTCCAAAC-3'	
	5' -GCTTGCAAAACACTGCAGTTGACTTGTGTTGGAG-3'	Primer for amplifying P4502C9 0.6Kb fragment
	5' -ACTGAGCAGCAGCCAGGCCATCTGCTCTTC-3'	
2E1	5' -CCCCAGAAATTCATGTCTGCCCTCGCAGTG-3'	Primer for amplifying P4502E1 0.5Kb fragment
	5' -CCTCTGGATCCGGCTCTCATTGCCCTGTTTC-3'	
	5' -GAAACAGGCCAATGACAGCCGGATCCAGAGG-3'	Primer for amplifying P4502E1 1.0Kb fragment
	5' -GAAACACTGTGTTGATCCGGGGGGTTCAGG-3'	

Fig. 1

3A4	5'-AGTAAGGAAATCTAGAAATGGCTCTCATCCACG-3'	Primer for amplifying P4503A4 0.6Kb fragment
	5'-ACGAGCTCCAGATCGGACAGAGCTTTG-3'	
	5'-CAAAAGCTCTGTCCGATCTGGAGCTCGT-3'	Primer for amplifying P4503A4 0.9Kb fragment
	5'-CAAAAGTAAATTCAGGTACCTGGTGTCTCAGCC-3'	
1A1	5'-CCTCTAGAAATGCTTTTCCCAATCTCCATG-3'	Primer for amplifying P4501A1 1.0Kb fragment
	5'-CCAATCACTGTGTGCGAGCTCCTCTTGGATC-3'	
	5'-GATCCAAGAGGAGCTCGACACAGTGATTGG-3'	Primer for amplifying P4501A1 0.5Kb fragment
	5'-GGGCTCTCAAGCACCTAAGAGCGCAGCTGC-3'	
2A6	5'-GCTTCTAGAAATGCTGGCCTCAGGGATGCTTC-3'	Primer for amplifying P4502A6 0.6Kb fragment
	5'-CGTGGAGGTTGACGTGAACCTGGAAGATTTC-3'	
	5'-GAATCTTCCAGTTCACGTCAACCTCCACG-3'	Primer for amplifying P4502A6 0.9Kb fragment
	5'-AGACCTGGTACCGCACAGCCCTCGCTCAG-3'	

Fig. 2

2B6	5' - CCTCTAGAAAAATGGAACTCAGCGTCCTCCT-3'	Primer for amplifying P4502B6 1.5Kb fragment
	5' - GGGGATCCTGAAATGACCCCTGGAAATCCTTTG-3'	
2C8	5' - GAAAGAGAAAGTCTAGAAATGGAAACCTTTTGTGGTCC-3'	Primer for amplifying P4502C8 1.5Kb fragment
	5' - ATAGCAGATCGGCAGCCAGATGGGCTAGCAATC-3'	
2C18	5' - AGTCTAGAAATGGTACCAGCTGTGGCTCTGG-3'	Primer for amplifying P4502C18 0.9Kb fragment
	5' - CCCCAAAACATATCAGTTACAGTGGCTATCAAGC-3'	
	5' - CCCGATTATTGGAAATATCCTGCAATTAGATG-3'	Primer for amplifying P4502C18 1.4Kb fragment
	5' - ACAGCACAGGAGCAGCCAAACTATCTGCC-3'	

Fig. 3

2019 The sequence shown by 5' - ... - 3' is described in SEQ
ID Nos: 20 to 40.

Primer for amplifying
P4502D6 0.4kb fragment

Primer for amplifying
P4502D6 0.9kb fragment

Primer for amplifying
P4503A4 XbaI-XhoI fragment

206 5'-TGTTTCAGCCCTGCAGCTGCCCTGGAC-3'
5'-AAGCGAGGGTCGTCGTATTCGAAAGCG-3'

5'-GCTTCGAAATACGACGACCCCTCGCTTCCCTC-3'
5'-ACTAGGTACCCCACTTCTAGCGGGGCACAG-3'

3A4 (An artificial fused enzyme)

5' - AATCTAGAAATCGCTCTCATCCAC - 3'
5' - AGGACTCGACCGCGCTCCACTTACCGTCCCATCC - 3'

Fig. 4

(1) Linker for cloning 1A2

5'-AGCTTAAAAAATGGCATTTGCCAGTGTGTTCCTTCTCGGCCACAGAGCT-3'
 3'-ATTTTITACCGTAACAGGGTCACACAAGGGAAGAGCCGGTGTG -5'

(2) Linker for cloning 2D6

5'-CTAGATATGGGGCTAGAACACTGGTGGCCCTTGGCGGTGATAGTGG-3'
 3'-TATACCCCGATCTTCGTGACCAACGGGGACCGGCACTATCACC-5'

5'-CCATCTTCCTGCTCCTGGTGGACCTGTATGCCACCGGGCCCAACGCTGGGCTGCCAGGCTACCCACAGGCCCCCTGGCCACTGCCCGGGGTGCA-3'
 3'-GGTAGAAGGACGAGGACCACCTGGACTACGTGGCCCGCGTTGGGACCCGACGTCGGATGGGTGGTCCGGGGACGGTGAACGGGCCG -5'

5'-GGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAAACACACCACTACTGCTTTCGACCAAGTTGGGGCGCCGCTTGGGGACGTTTCAGGCTGCA-3'
 3'-CCCGACCCGTTGGACGACGTACACCTGAGGGTCTTGTGTGGTATGACGAAGCTGGTCAACGCCCGCGGGCAAGCCCCCTGCAAGTCCG -5'

Fig. 5

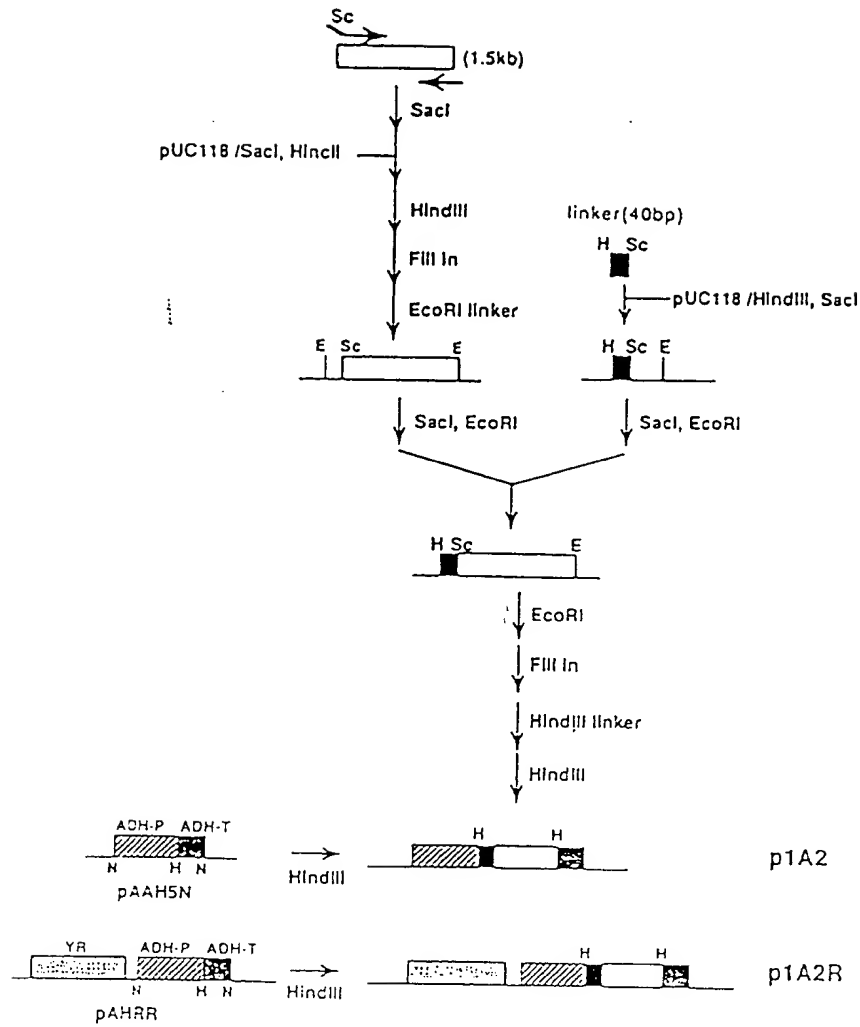


Fig. 6

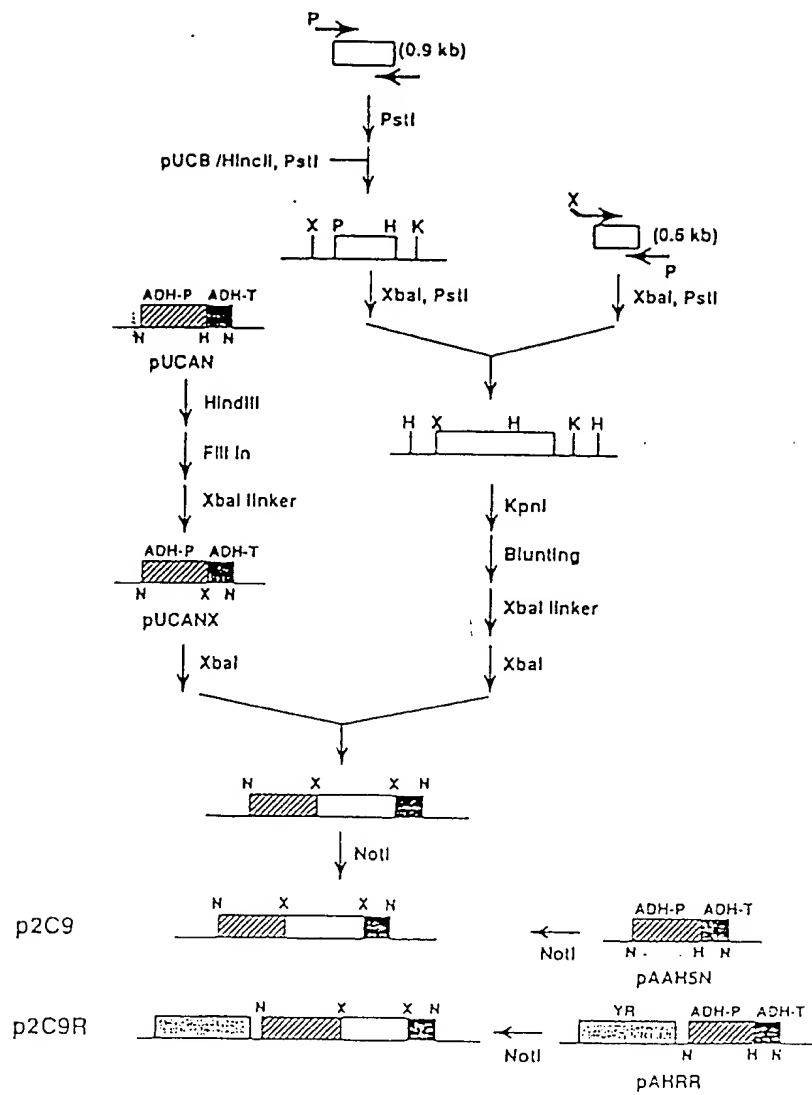


Fig. 7

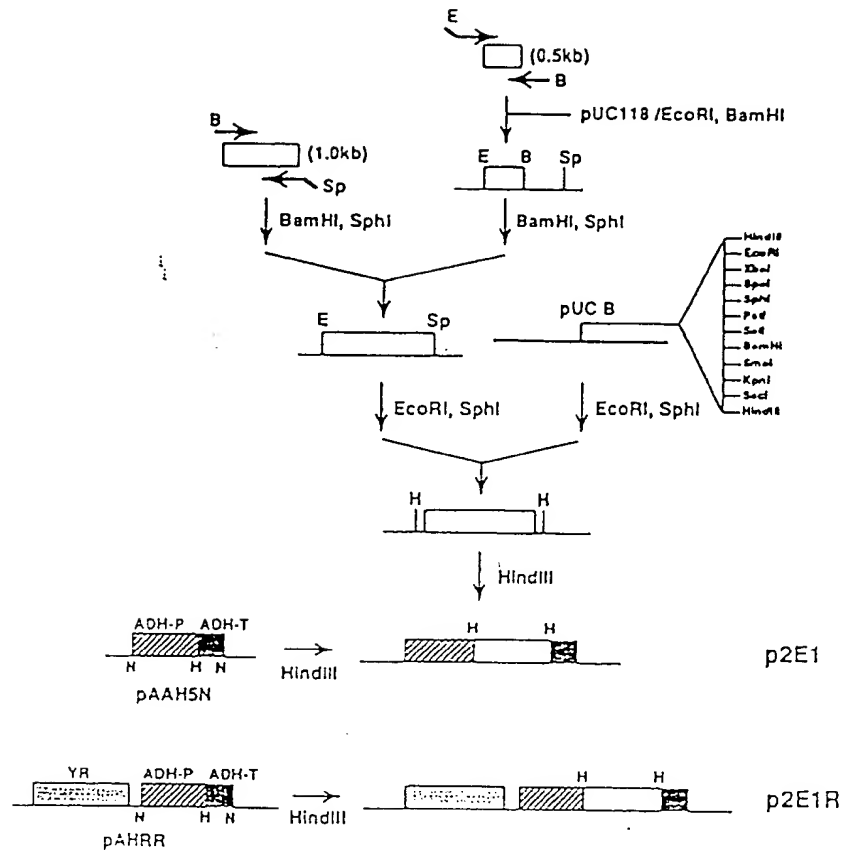


Fig. 8

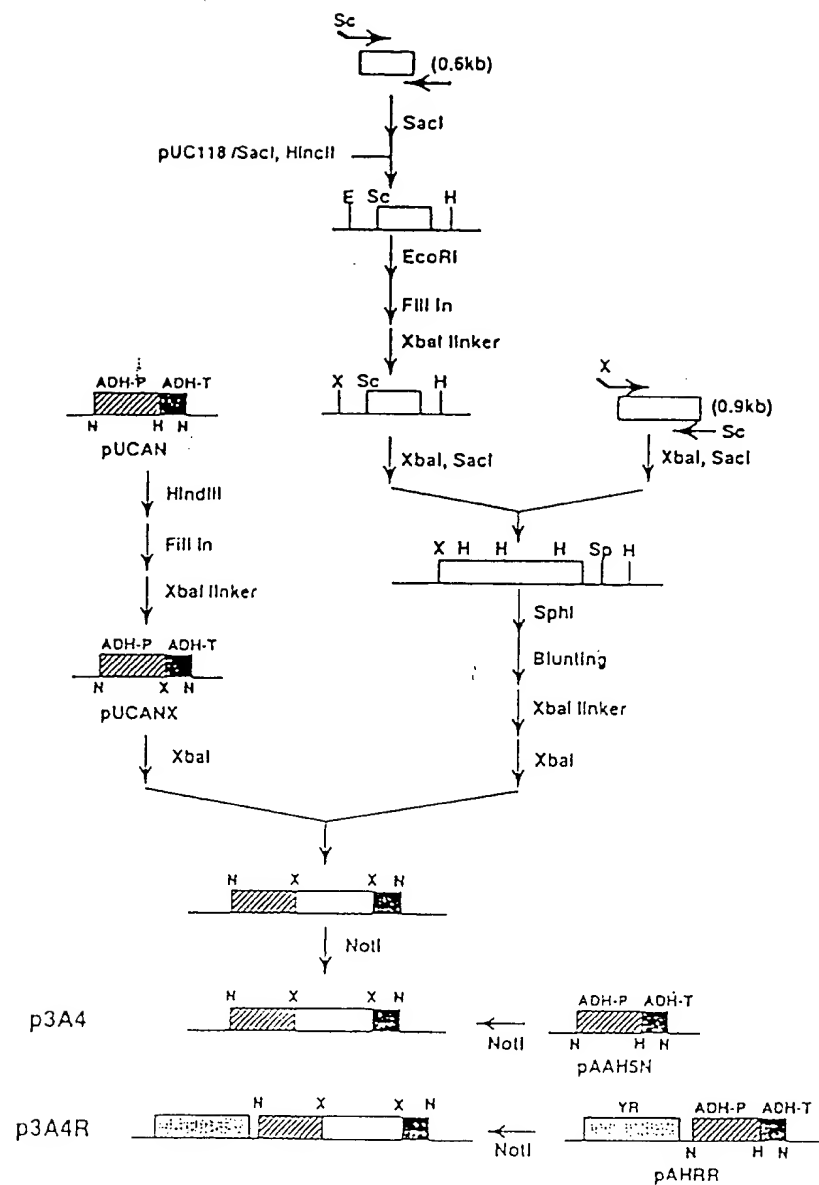


Fig. 9

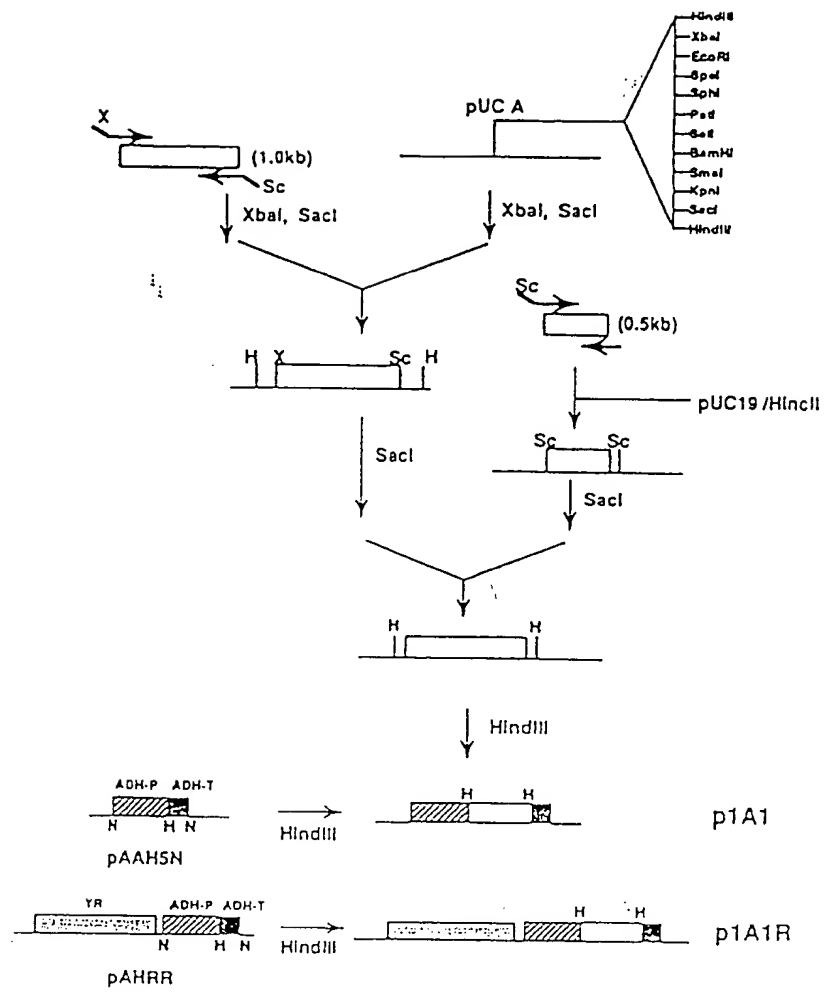


Fig. 10

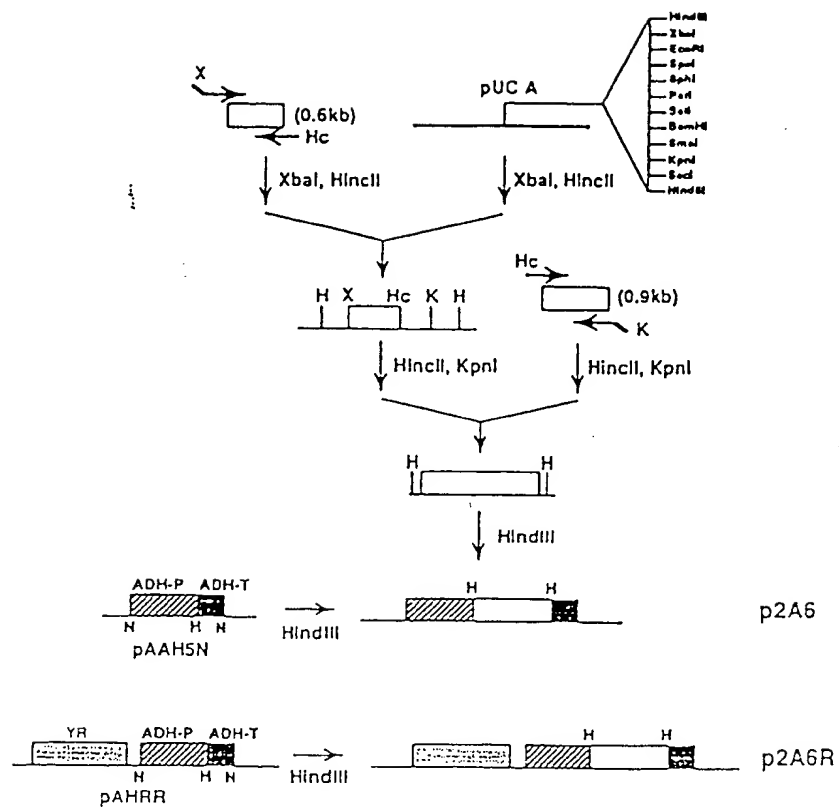


Fig. 11

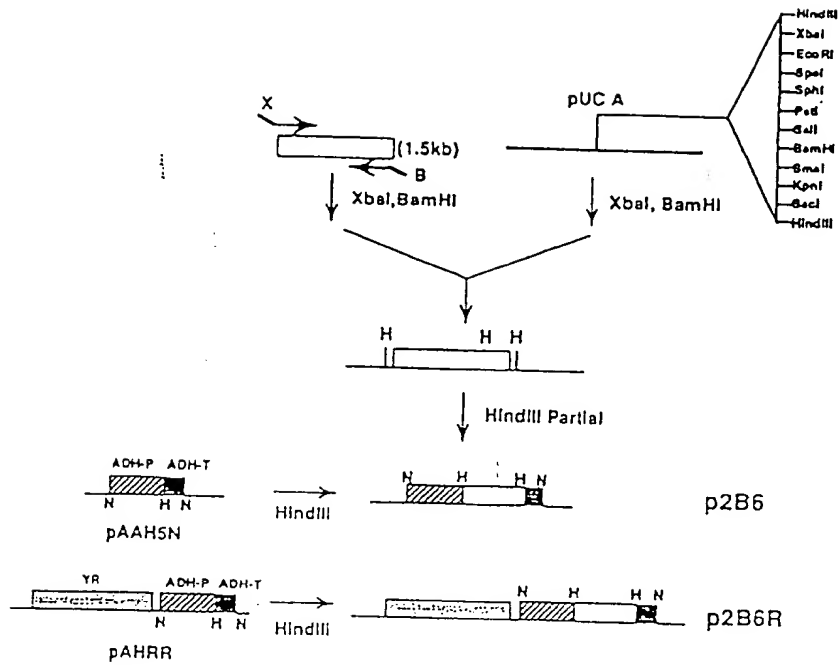


Fig. 12

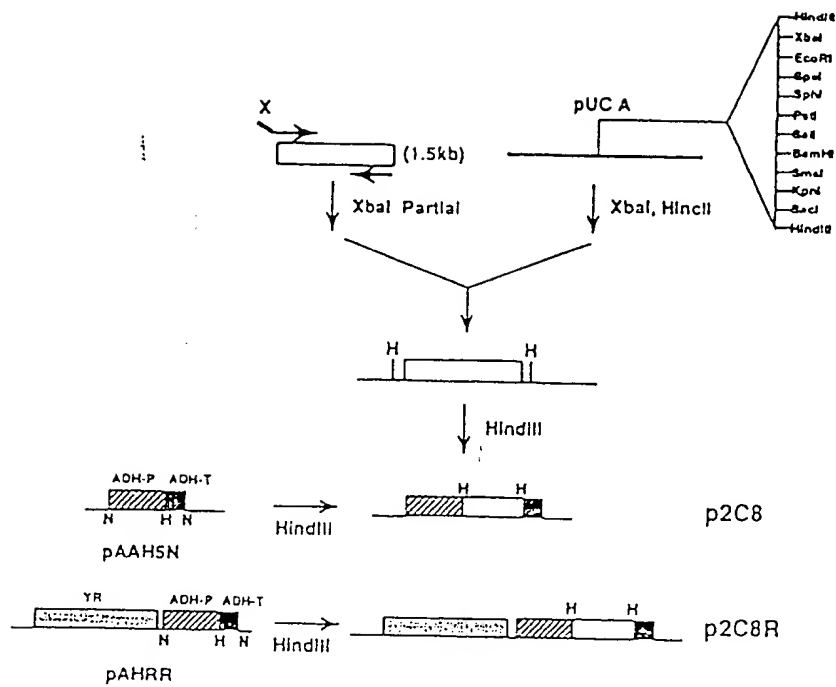


Fig. 13

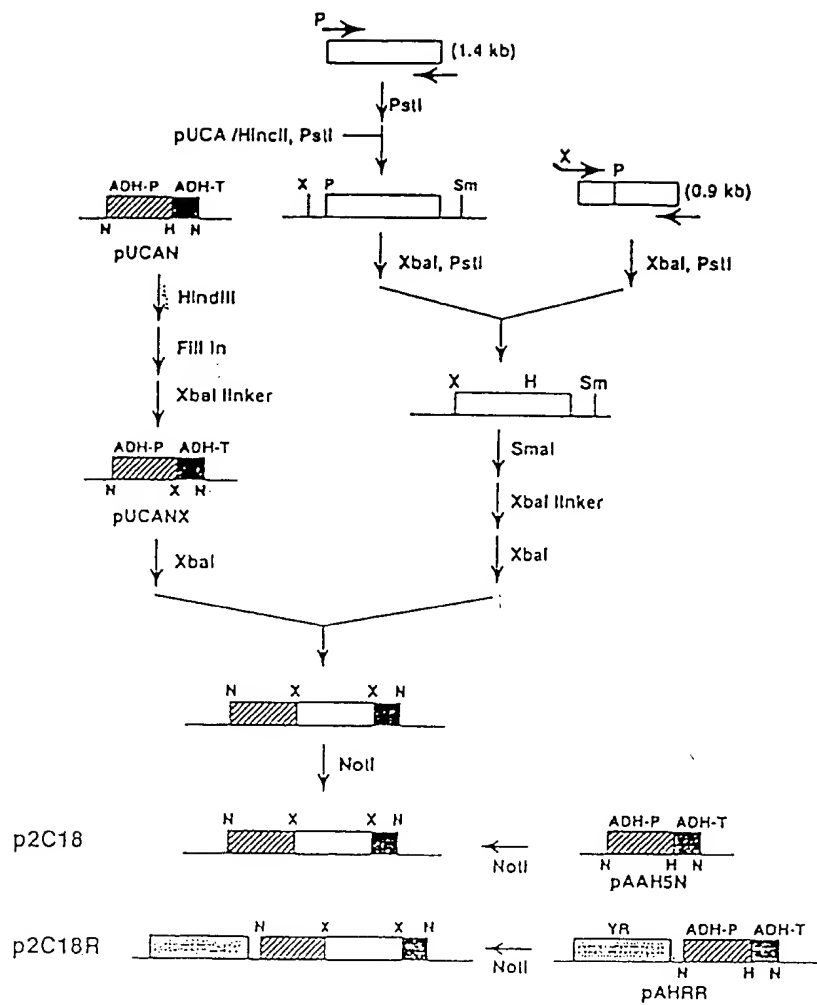
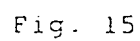
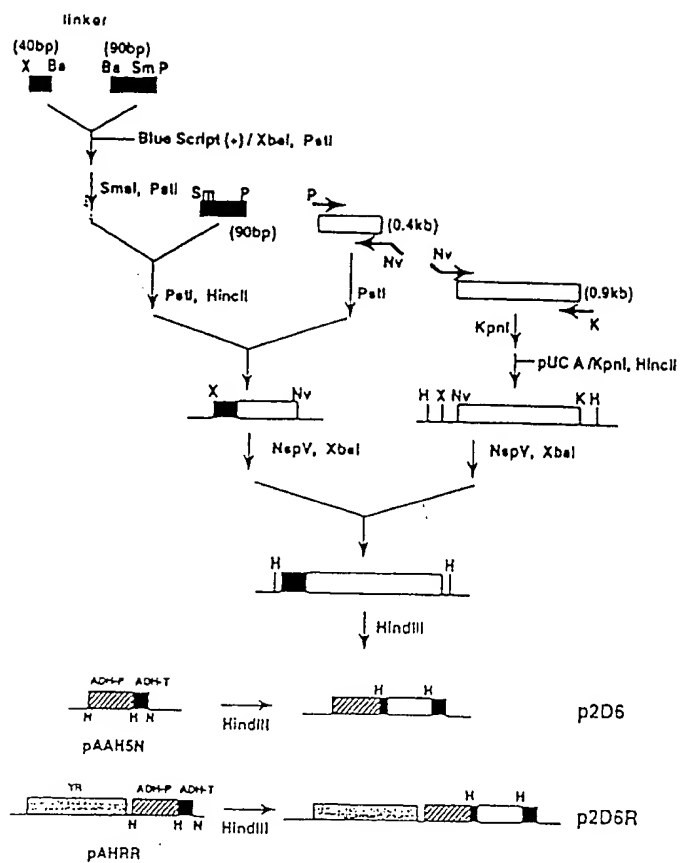


Fig. 14





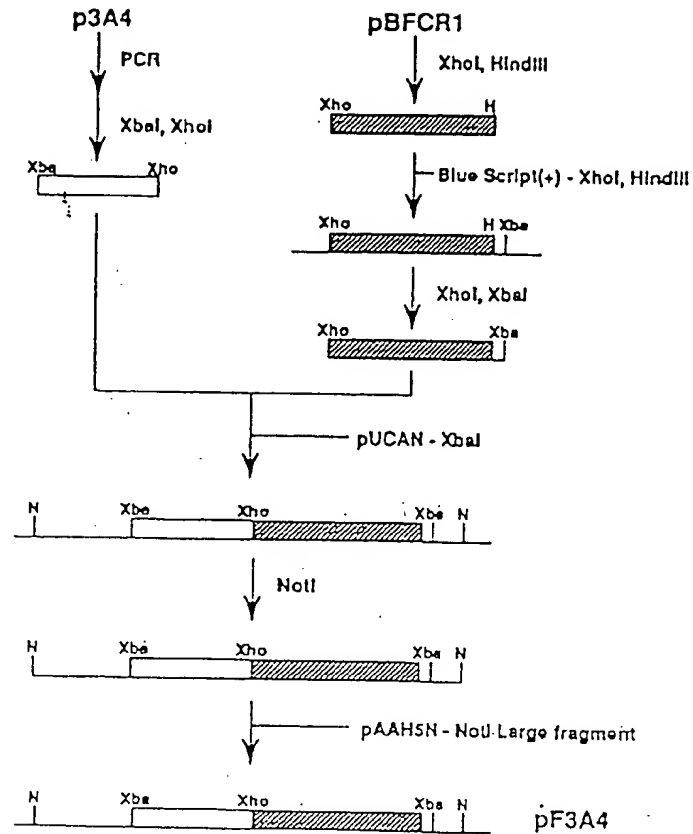


Fig. 17

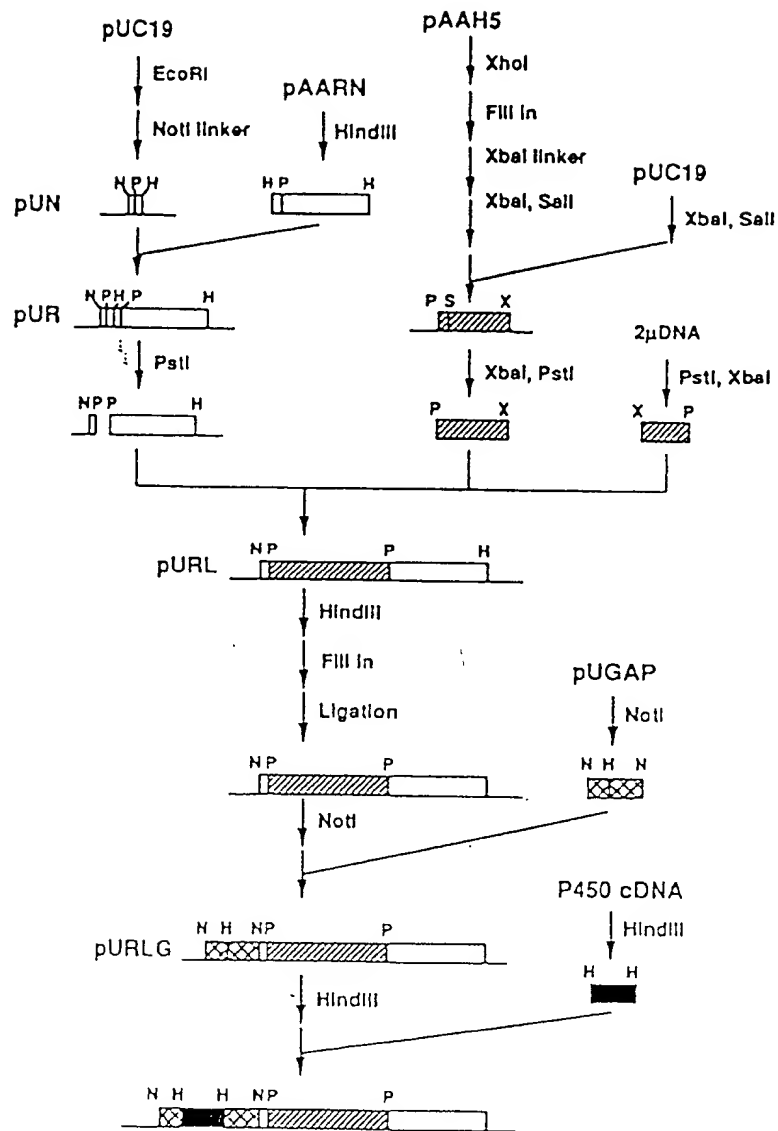


Fig. 18